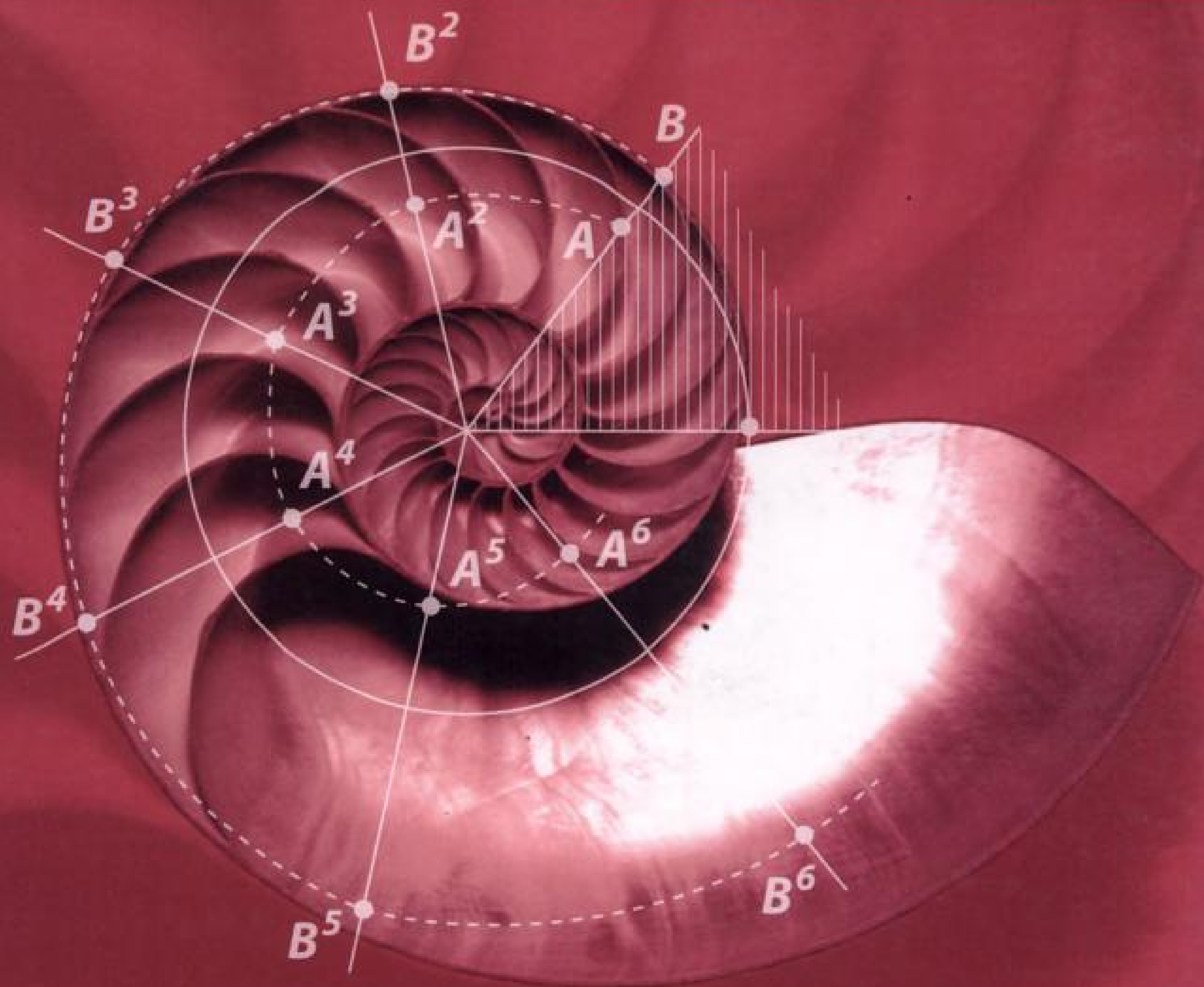
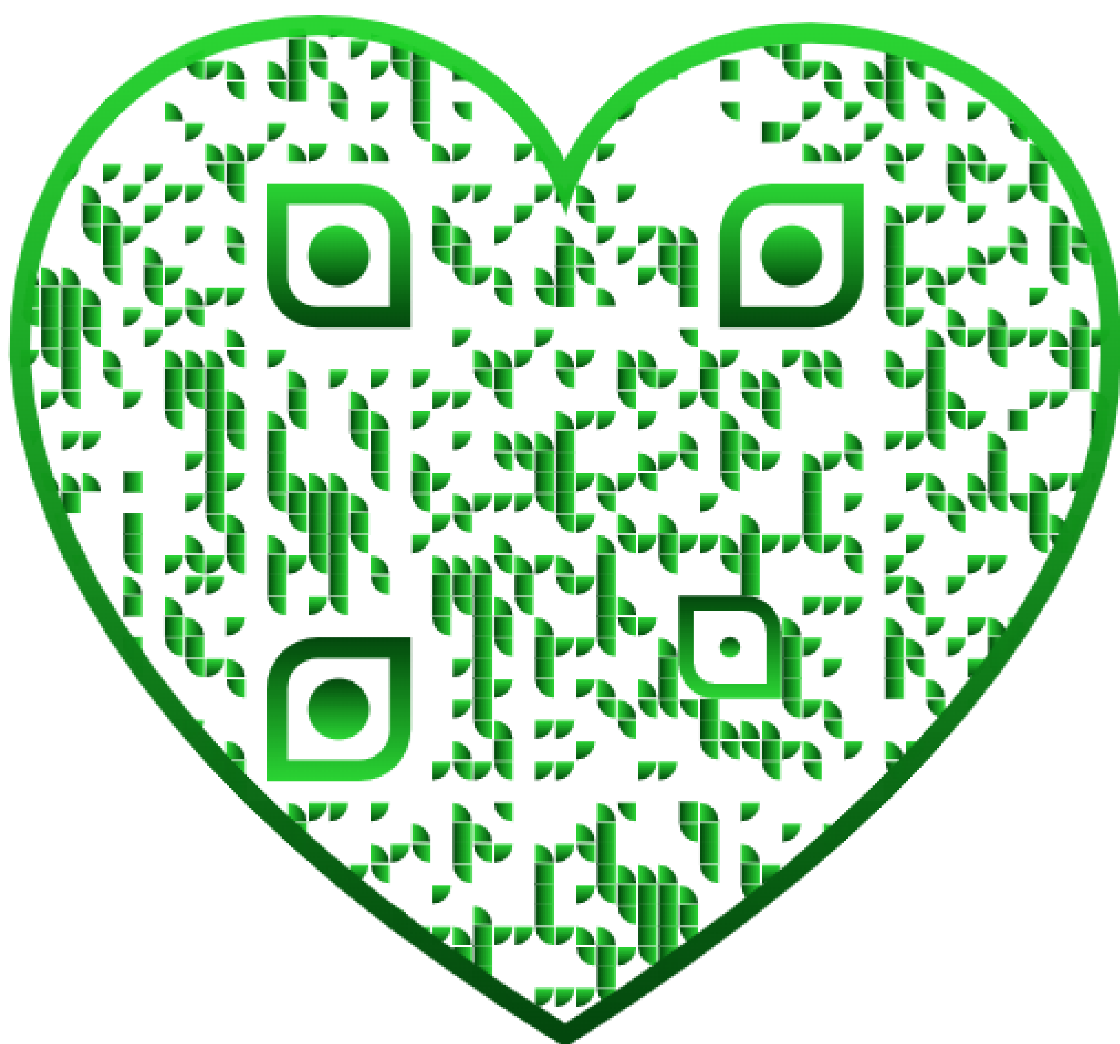
THOMSON



ÁLGEBRA LINEAR

David Poole

Material com direitos autorais



THOMSON

Gerente Editorial:
Adilson Pereira

Editor de Desenvolvimento:
Marcio Coelho

Produtora Editorial:
Ada Santos Seles

Produtora Gráfica:
Patricia La Rosa

Título Original:
Linear algebra: a modern introduction
ISBN: 0-534-34174-8

Tradutoras Técnicas:
Martha Salerno Monteiro (Coord.),
Célia Mendes Carvalho Lopes,
Fernanda Soares Pinto Cardona,
Iole de Freitas Druck,
Leila Maria Vasconcellos Figueiredo,
Maria Lúcia Sobral Singer e
Zara Issa Abud

Copidesque:
Glauco Peres Damas

Revisão:
Danae Stephan e Marcos Soel

Editoração Eletrônica:
Know-how Editorial

Capa:
Megaart Design

COPYRIGHT © 2003 de Brooks/Cole
COPYRIGHT © 2004 para a língua portuguesa adquirido por Pioneira Thomson Learning Ltda., uma divisão da Thomson Learning, Inc.
Thomson Learning™ é uma marca registrada aqui utilizada sob licença.

Impresso no Brasil.
Printed in Brazil.
1 2 3 4 06 05 04

Rua Traipu, 114 – 3º andar
Perdizes – CEP 01235-000
São Paulo – SP
Tel.: (11) 3665-9900
Fax: (11) 3665-9901
sac@thomsonlearning.com.br
www.thomsonlearning.com.br

Todos os direitos reservados. Nenhuma parte deste livro poderá ser reproduzida, sejam quais forem os meios empregados, sem a permissão, por escrito, da Editora. Aos infratores aplicam-se as sanções previstas nos artigos 102, 104, 106 e 107 da Lei nº 9.610, de 19 de fevereiro de 1998.

Dados Internacionais de Catalogação na Publicação (CIP)
(Câmara Brasileira do Livro, SP, Brasil)
Poole, David.
Álgebra linear / David Poole, 1955 ; tradutoras técnicas Martha Salerno Monteiro [(coord.)... [et al.] --; São Paulo : Pioneira Thomson Learning, 2004.
Título original: Linear algebra : a modern introduction
Várias tradutoras.
ISBN 85-221-0359-3.
Álgebra linear I. Título.
03-5647 CDD-512.5
Índices para catálogo sistemático:
1. Álgebra linear : Matemática 512.5

Material com direitos autorais

◆ Capítulo 3 Matrizes

3.1 Introdução: Matrizes em Ação *125*
3.2 Operações com Matrizes *127*
3.3 A Álgebra de Matrizes *142*
3.4 A Inversa de uma Matriz *151*
3.5 Subespaços, Base, Dimensão e Posto *167*
3.6 Introdução às Transformações Lineares *187*
3.7 Aplicações *201*
Investigação: A Fatoração LU 225

◆ Capítulo 4 Valores e Vetores Próprios

4.1 Introdução: Um Sistema Dinâmico sobre Grafos *230*
4.2 Introdução aos Valores e Vetores Próprios *232*
4.3 Determinantes *239*
Investigação: Aplicações Geométricas de Determinantes 259
4.4 Valores e Vetores Próprios de Matrizes $n \times n$ *264*
4.5 Semelhança e Diagonalização *272*
4.6 Métodos Iterativos para o Cálculo de Autovalores *282*
4.7 Aplicações e o Teorema de Perron-Frobenius *291*

◆ Capítulo 5 Ortogonalidade

5.1 Introdução: Sombras em uma Parede *320*
5.2 Ortogonalidade em $\mathbb{R}^n$ *323*
5.3 Complementos e Projeções Ortogonais *332*
5.4 O Processo de Gram-Schmidt e a Fatoração *QR* *342*
Investigação: A Fatoração QR Modificada 350
Investigação: Aproximação de Autovalores Usando Algoritmo QR 352
5.5 Diagonalização de Matrizes Simétricas *353*
5.6 Aplicações *361*

◆ Capítulo 6 Espaços Vetoriais

6.1 Introdução: Quadrados Mágicos *385*
6.2 Espaços Vetoriais e Subespaços *387*
6.3 Dependência Linear, Base e Dimensão *401*
6.4 Mudança de Base *415*
6.5 Transformações Lineares *424*
6.6 O Núcleo e a Imagem de uma Transformação Linear *433*
6.7 A Matriz de uma Transformação Linear *447*
Investigação: Mosaicos, Reticulados e a Restrição Cristalográfica 465
6.8 Aplicações *468*

Material com direitos autorais

◆ **Capítulo 7** Distância e Aproximação

7.1 Introdução: A Geometria do Taxista *484*
7.2 Espaços com Produto Interno *486*
Investigação: Desigualdades Geométricas e Problemas de Otimização *498*
7.3 Normas e Funções Distância *502*
7.4 Método dos Mínimos Quadrados *518*
7.5 A Decomposição por Valores Singulares *540*
7.6 Aplicações *559*

◆ **Apêndice A** A Notação Matemática e Métodos de Demonstração *572*

◆ **Apêndice B** Indução Matemática *581*

◆ **Apêndice C** Números Complexos *588*

◆ **Apêndice D** Polinômios *600*

◆ **Apêndice E** Bytes de Tecnologia *610*

Respostas a Exercícios Selecionados de Número Ímpar *652*

Índice Remissivo *I.1*

Material com direitos autorais

Prefácio

A última coisa que se sabe quando se escreve
um livro é o que colocar primeiro.

Blaise Pascal
Pensées, 1670

Este livro foi projetado para uso em uma disciplina introdutória de álgebra linear, com duração de um ou dois semestres, como ela é tipicamente cursada por alunos de primeiro ou segundo ano. Escrevi o livro com esse público em mente. Ele é, antes de tudo, destinado aos alunos: fiz o melhor que pude para escrevê-lo de modo que os alunos o achassem fácil de ler e, além disso, tivessem vontade de lê-lo. (Claro que ficarei satisfeito se os professores sentirem o mesmo!)

Uma disciplina introdutória de álgebra linear costuma ter alunos provenientes de vários cursos de graduação. Além de alunos de cursos de Matemática, é possível haver alunos vindos de cursos de Engenharia, Física, Química, Ciência da Computação, Biologia, Ciência Ambiental, Geografia, Economia, Psicologia, Administração e Educação, bem como outros alunos que cursam a disciplina como eletiva ou para cumprir requisitos para a graduação. Destinar um livro a um grupo de estudantes potencialmente tão diverso apresenta vários desafios: como balancear teorias e aplicações? A exposição deve ser rigorosa ou informal? A abordagem expositiva tradicional é incompatível com um aprendizado centrado no estudante? Qual o papel da tecnologia e das considerações computacionais relacionadas com ela? Pode a álgebra linear ser tornada "enxuta e viva"?[*, 1] Neste Prefácio, descreverei como essas questões foram abordadas (de modo bem-sucedido, espero) em meu texto.

Não existe um único estilo de aprendizado. Em qualquer classe haverá alguns alunos que trabalham bem individualmente e outros que trabalham melhor em grupos, alguns que preferem o aprendizado com base em leituras e outros que prosperam em um ambiente de oficina didática, alguns que apreciam manipulações algébricas, outros adeptos de métodos numéricos (com ou sem um computador) e alguns que exibem forte intuição geométrica. Um bom livro-texto usa uma variedade de maneiras para apresentar o material, para que todos os tipos de estudante possam encontrar um caminho a seguir. Em consonância com essas idéias, apresento os tópicos algebricamente, geometricamente, numericamente e verbalmente.

Este livro reflete a minha crença – baseada em aproximadamente 20 anos de experiência com ensino – no fato de as técnicas pedagógicas apropriadas para alunos de outros cursos de graduação serem igualmente apropriadas para alunos de cursos de graduação em Matemática. Estratégias de ensino que incorporam pro-

* N.T.: *Lean and lively*, no original em inglês.

1 Esta expressão é identificada como "movimento de reforma" do cálculo, movimento este originado com a obra *Toward a lean and lively calculus*, de R. G. Douglas (Washington, DC: Mathematical Association of America, 1986). O mesmo sentimento é igualmente aplicável à álgebra linear.

Material com direitos autorais

jetos de estudo, experiências de aprendizado em grupo, trabalhos escritos e o uso de tecnologia são tão apropriadas em uma classe de bons alunos de Matemática como em uma disciplina de serviço. Uma disciplina de álgebra linear voltada à preparação dos alunos para disciplinas teóricas de nível mais avançado deve também expor esses alunos às aplicações. Reciprocamente, uma disciplina de álgebra linear aplicada ou numérica deve mencionar alguns dos resultados teóricos que justificam as técnicas. E todos os estudantes de álgebra linear devem pelo menos ganhar consciência dos métodos numéricos empregados na prática para a resolução, com dados reais, de problemas em grande escala.

A abordagem deste livro é compatível com a maioria das recomendações feitas pelo Grupo de Estudos Curriculares de Álgebra Linear para uma primeira disciplina de álgebra linear.[2] O ímpeto geral dessas recomendações, baseadas em uma oficina realizada em 1990 no College of William and Mary, é se afastar da abstração em direção a uma disciplina mais concreta, baseada em matrizes. Do ponto de vista pedagógico, não há dúvida de que, para a maioria dos alunos, exemplos concretos devem preceder a abstração. Segui essa abordagem aqui. Entretanto, não acredito que matrizes e sistemas de equações lineares devam aparecer primeiro. Álgebra linear é essencialmente o estudo sobre vetores e, portanto, os alunos devem ver vetores antes (em uma situação concreta) para ganhar alguma intuição geométrica. Além disso, apresentar com antecedência os vetores permite que os alunos vejam como os sistemas de equações lineares surgem naturalmente de problemas geométricos. Matrizes surgem, então, de modo igualmente natural, dos sistemas lineares, como matrizes de coeficientes e como agentes de mudança (transformações lineares). Pode-se dizer que este livro segue uma abordagem baseada em vetores, na qual as matrizes desempenham um importante papel de apoio.

Em consonância com a temática "enxuta e viva", tentei limitar o número de teoremas no texto. Na maioria dos casos, os resultados rotulados como teoremas são utilizados mais tarde no texto ou resumem um trabalho precedente. Resultados interessantes, porém não centrais para o livro, foram incluídos como exercícios ou investigações. Um exemplo é o produto vetorial de dois vetores, tratado somente em investigações (nos Capítulos 1 e 4). Diferentemente de muitos livros, este não tem um capítulo sobre determinantes. Os resultados essenciais estão todos na Seção 4.3, e o material adicional de interesse está na investigação "Aplicações Geométricas dos Determinantes". O livro é, no entanto, abrangente para um texto introdutório. Sempre que possível, incluí demonstrações elementares e acessíveis de teoremas, para evitar a necessidade de dizer que "a demonstração deste resultado está além do alcance deste texto". O resultado é, espero, uma obra autocontida.

Não fui parcimonioso com as aplicações. A quantidade delas no livro excede bastante o que se pode cobrir em uma só disciplina. Entretanto, é importante que os alunos vejam a extensão impressionante de problemas aos quais a álgebra linear pode ser aplicada. Incluí também material moderno sobre teoria da codificação, assunto não encontrado normalmente em textos introdutórios de álgebra linear.

Álgebra linear é um assunto estimulante. Como ponto de entrada, conduz aos muitos ramos da matemática com os quais tem ligações, incluindo álgebra abstrata, matemática discreta, cálculo, equações diferenciais, geometria, estatística, métodos numéricos, pesquisa operacional e sistemas dinâmicos. Ela expõe os estudantes a aspectos teóricos, aspectos aplicados e aspectos numéricos da matemática e tem aplicações em uma gama variada de disciplinas. Espero que, com este livro, os estudantes aprendam a valorizar a beleza, o poder e a imensa utilidade da álgebra linear, e que o façam se divertindo.

Características

Estilo de Escrita Claro

O texto é escrito em um estilo simples, direto e coloquial. Tanto quanto possível, usei "português matemático" em vez de me basear em notação matemática. Todas as demonstrações apresentadas, porém, são

[2] Veja *The College Mathematics Journal*, n. 24, 1993, p. 41-46.

Material com direitos autorais

completamente rigorosas, e o Apêndice A contém uma introdução à notação matemática para aqueles que desejarem empregá-la em seu próprio texto. Exemplos concretos aparecem quase sempre antes dos teoremas, que são então seguidos por exemplos adicionais e aplicações. Esse percurso – do específico para o geral e de volta ao específico – é consistentemente realizado ao longo do livro.

Conceitos-Chave Apresentados com Antecedência

Muitos alunos encontram dificuldade quando uma disciplina de Álgebra Linear vai da parte prática (resolução de sistemas de equações lineares, manipulação de vetores e matrizes) para a parte teórica (conjuntos geradores, independência linear, subespaços, bases e dimensão). Este livro apresenta com antecedência todos os conceitos-chave da álgebra linear, em um cenário concreto, antes de revisitá-los com generalidade total. Conceitos vetoriais como produto escalar, comprimento, ortogonalidade e projeção são discutidos primeiro no Capítulo 1, no cenário concreto do $\mathbb{R}^2$ e $\mathbb{R}^3$, antes de as noções mais gerais de produto interno, norma e projeção ortogonal aparecerem nos Capítulos 5 e 7. Analogamente, as noções de conjunto gerador e independência linear recebem um tratamento concreto no Capítulo 2, antes de serem generalizadas para espaços vetoriais no Capítulo 6. Os conceitos fundamentais de subespaço, base e dimensão aparecem primeiro no Capítulo 3, quando são apresentados o espaço linha, o espaço coluna e o espaço nulo de uma matriz, mas somente no Capítulo 6 essas idéias recebem um tratamento geral. Autovetores e autovalores são apresentados e explorados no Capítulo 4, para matrizes 2×2, antes de aparecerem no caso $n\times n$. No início do Capítulo 4, teremos apresentado todos os conceitos-chave da álgebra linear, com exemplos numéricos completos para lhes dar apoio. Quando, mais adiante no livro, essas idéias aparecerem com toda a generalidade, os alunos já poderão estar acostumados com elas e por isso não se sentirão muito intimidados.

Ênfase em Vetores e Geometria

Seguindo a filosofia de que álgebra linear é, fundamentalmente, o estudo de vetores, este livro salienta a intuição geométrica. Em conformidade com isso, o primeiro capítulo é sobre vetores e desenvolve muitos conceitos que aparecem repetidamente ao longo do texto. Conceitos como ortogonalidade, projeção e combinação linear são todos encontrados no Capítulo 1, o qual contém ainda um tratamento abrangente do estudo de retas e de planos em $\mathbb{R}^3$, o qual traz idéias essenciais para a resolução de sistemas de equações lineares. Essa ênfase em vetores, geometria e visualização é recorrente ao longo do texto. O Capítulo 4 apresenta autovetores usando "autofiguras" como apoio visual. A demonstração do Teorema de Perron é vista heuristicamente em um primeiro momento para depois ser vista formalmente, com o emprego de um argumento geométrico nos dois casos. A natureza da decomposição por valores singulares é também explicada informalmente com o uso de um argumento geométrico.

Investigações

Cada capítulo começa com uma investigação dirigida (Seção 1), na qual os alunos são convidados a descobrir – individualmente ou em grupos – algum aspecto do capítulo que ela inicia. Exemplificando: "O Jogo da Pista de Corrida" apresenta vetores, "Matrizes em Ação" mostra a multiplicação de matrizes e as transformações lineares, "Quadrados Mágicos" refere-se aos espaços vetoriais e "A Geometria dos Taxistas" prepara para a generalização dos conceitos de norma e de distância. Investigações adicionais encontradas ao longo do livro incluem aplicações de vetores e determinantes à geometria, um estudo de simetria por meio dos mosaicos de M. C. Escher e problemas de otimização que empregam desigualdades geométricas. Há também investigações que apresentam importantes considerações numéricas, análise de algoritmos e fatoração LU. Atribuir algumas dessas investigações como tarefas aos alunos é uma forma de encorajá-los a se tornarem aprendizes ativos e de torná-los "donos" de uma pequena parte da disciplina.

Material com direitos autorais

Aplicações

O livro contém uma seleção abundante de aplicações escolhidas de uma ampla gama de áreas, incluindo matemática, ciência da computação, física, química, engenharia, biologia, psicologia, geografia e sociologia. Dentre essas aplicações, é digno de nota o forte tratamento da teoria da codificação, desde os códigos detectores de erros (tais como o International Standard Book Number, ou ISBN) até os sofisticados códigos corretores de erros (como o código de Reed-Muller, empregado para transmitir, do espaço, as fotografias feitas por satélites).

Exemplos e Exercícios

Há aproximadamente 400 exemplos neste livro, a maior parte deles com mais detalhes do que é de costume em um livro texto-introdutório de álgebra linear. Esse nível de detalhe está de acordo com a filosofia de que os alunos devem desejar (e ser capazes de) ler um livro-texto. Não se espera que todos esses exemplos sejam vistos em classe; muitos podem ser designados para estudo individual ou em grupo, possivelmente como parte de um projeto. A maioria dos exemplos tem como contrapartida pelo menos um exercício, para que os alunos possam testar as aptidões cobertas no exemplo antes de explorar generalizações.

O livro inclui mais exercícios (acima de 1.800) do que a maioria dos livros-textos em um nível similar. As respostas para a maior parte dos exercícios práticos de número ímpar podem ser encontradas ao final do livro. Isso significa que os alunos podem estudar um exemplo no texto, praticar com um exercício com número ímpar e conferir a resposta antes de trabalhar em um exercício proposto. Professores encontrarão uma abundância de exercícios, dentre os quais poderão selecionar tarefas propostas. Os exercícios em cada seção aparecem em uma progressão gradativa, desde os rotineiros até os desafiadores. Eles variam dos voltados para o cálculo manual aos que requerem o uso de uma calculadora ou de um sistema de álgebra computadorizado. Variam ainda de exercícios teóricos e numéricos a exercícios conceituais. Muitos dos exemplos e exercícios usam dados reais compilados de situações verídicas. Há, por exemplo, problemas sobre modelagem do crescimento das populações de renas e focas, avaliação da data do monumento de Stonehenge usando carbono radiativo e predição dos salários de jogadores profissionais de beisebol. O trabalho nesses problemas reforça o fato de a álgebra linear ser uma ferramenta valiosa para a modelagem de problemas da vida real.

Resumos Biográficos e Notas Etimológicas

É importante que os estudantes aprendam alguma coisa sobre a história da matemática e que venham a percebê-la como um esforço tanto social e cultural quanto científico. Assim, o texto contém pequenos esboços biográficos de muitos dos matemáticos que contribuíram para o desenvolvimento da álgebra linear. Espero que eles ajudem a colocar uma face humana no assunto e a proporcionar aos estudantes um outro modo de se relacionar com a matéria.

Percebi que muitos estudantes sentem-se alienados da matemática porque a terminologia não faz sentido para eles – ela é simplesmente uma coleção de palavras a serem aprendidas. Para ajudar a superar esse problema, incluí pequenas notas etimológicas que dão a origem de muitos termos usados em álgebra linear. (Por exemplo, por que usamos a palavra "normal" para nos referirmos a um vetor perpendicular a um plano?)

Ícones de Margem

As margens do livro contêm vários ícones cujo objetivo é chamar a atenção do leitor de diversas maneiras. Cálculo diferencial e integral não é um pré-requisito para este livro, mas a álgebra linear tem muitas aplicações interessantes e importantes no cálculo. O ícone $\frac{dy}{dx}$ denota um exemplo ou exercício que requer cálculo.

Material com direitos autorais

(Esses exercícios podem ser omitidos se nem todos na classe tiveram pelo menos um semestre de cálculo. Alternativamente, eles podem ser propostos como projetos.) O ícone a + bi denota um exemplo ou exercício que envolve números complexos. (Para estudantes que não têm familiaridade com números complexos, o Apêndice C contém todo o material básico necessário.) O ícone CAS indica que é necessário – ou pelo menos útil – um sistema algébrico computacional (por exemplo, Maple, Mathematica ou MATLAB) ou uma calculadora com capacidade de lidar com matrizes (como TI-89, TI-92 ou HP-48G) para a resolução do exemplo ou exercício.

Em um esforço para ajudar os alunos a aprender como ler e usar este livro-texto mais eficientemente, assinalei vários locais nos quais o leitor é aconselhado a fazer uma pausa. Esses podem ser lugares em que uma conta seja necessária, parte de uma demonstração deva ser complementada, uma afirmação deva ser verificada ou algum pensamento extra seja requerido. Um ícone com uma flecha aparece na margem em tais locais; a mensagem é "Vá mais devagar. Pegue seu lápis. Pense sobre isto".

Tecnologia

Este livro pode ser usado com sucesso por estudantes, tendo eles ou não acesso à tecnologia. Contudo, calculadoras com capacidade de efetuar operações com matrizes e sistemas algébricos computacionais (CAS) são atualmente acessíveis e, se usadas de maneira apropriada, podem enriquecer a experiência do aprendizado, bem como ajudar com os cálculos tediosos. Neste texto, adoto o ponto de vista de que os alunos precisam dominar todas as técnicas básicas de álgebra linear resolvendo à mão exemplos que não envolvam contas difíceis. A tecnologia pode então ser usada para resolver exemplos subseqüentes e aplicações, e para aplicar técnicas que se apóiem nas anteriores. Por exemplo, quando são introduzidos sistemas de equações lineares, são apresentadas soluções detalhadas; mais tarde, as soluções são simplesmente citadas e espera-se que o leitor as confira. Este é um bom momento para usar alguma forma de tecnologia. Da mesma maneira, quando aplicações usarem dados que tornem os cálculos à mão não práticos, use tecnologia. Todos os métodos numéricos que são discutidos dependem do uso de tecnologia.

Com ajuda da tecnologia, os estudantes podem explorar a álgebra linear de algumas maneiras estimulantes e descobrir muito por si mesmos. Por exemplo, se um dos coeficientes de um sistema linear for trocado por um parâmetro, quanto irão variar as soluções? Quanto a mudança de um único elemento de uma matriz afetará seus autovalores? Este livro não é um manual de tecnologia, e, em partes em que a tecnologia pode ser usada, não especifiquei um tipo particular de tecnologia. O Apêndice E, chamado "Bytes de Tecnologia", dá instruções para resolver uma seleção de exemplos de cada capítulo usando Maple, Mathematica e MATLAB. Imitando esses exemplos, os alunos poderão fazer mais cálculos e explorações usando qualquer um dos CAS que eles tiverem.

Álgebra Linear Numérica e Álgebra Linear Finita

O texto aborda dois aspectos de álgebra linear que são raramente mencionados juntos: álgebra linear finita e álgebra linear numérica. Com a introdução, logo no início, de aritmética modular, fui capaz de fazer da álgebra linear finita (mais apropriadamente, "álgebra linear sobre corpos finitos", embora eu não use essa frase) um tema recorrente ao longo do livro. Esse enfoque fornece acesso ao material sobre teoria de códigos nas Seções 1.5, 3.7, 5.6, 6.8 e 7.6. Há também uma aplicação a jogos lineares finitos, na Seção 2.5, que os estudantes realmente apreciam. Além de serem expostos às aplicações de álgebra linear finita, alunos de cursos de Matemática irão beneficiar-se ao estudar álgebra linear sobre corpos finitos, já que irão encontrá-los em outros cursos, tais como Matemática Discreta, Álgebra Abstrata e Teoria dos Números.

Todos os estudantes devem estar cientes de que, na prática, é impossível chegar a soluções exatas de problemas de larga escala em álgebra linear. A exposição a algumas das técnicas de álgebra linear numérica proporcionará

Material com direitos autorais

uma indicação de como obter soluções aproximadas com grande precisão. Alguns dos tópicos numéricos incluídos no livro são: erros de arredondamento e pivoteamento parcial, métodos iterativos de resolução de sistemas lineares e cálculo de autovalores, as fatorações *LU* e *QR* de matrizes, normas de matrizes, método dos mínimos quadrados e a decomposição por valores singulares. A inclusão de álgebra linear numérica também traz questões interessantes e importantes que são completamente ausentes da *teoria* da álgebra linear, tais como estratégias para pivoteamento, a condição de um sistema linear e a convergência dos métodos iterativos. Este livro não apenas levanta essas questões como também mostra como se pode abordá-las. As normas de matrizes e os valores singulares de uma matriz, ambos discutidos no Capítulo 7, são úteis a esse respeito.

Apêndices

O Apêndice A contém uma visão geral da notação matemática e dos métodos de demonstração; o Apêndice B discute a indução matemática. Todos os estudantes irão beneficiar-se da leitura dessas seções, mas alunos de cursos com ênfase em matemática poderão desejar prestar uma atenção especial neles. Alguns dos exemplos nesses apêndices são raros (por exemplo, o Exemplo 6 do Apêndice B) e enfatizam o poder dos métodos. O Apêndice C é uma introdução aos números complexos. Para alunos familiarizados com esses resultados, ele pode servir como uma referência útil; para outros, a seção terá todo o material de que eles necessitam para a compreensão das partes do texto que usam números complexos. O Apêndice D é sobre polinômios. Percebi que muitos estudantes precisam de uma revisão sobre esses fatos. A maioria dos estudantes não terá familiaridade com a Regra de Sinais de Descartes; ela é usada no Capítulo 4 para explicar o comportamento dos autovalores das matrizes de Leslie. O Apêndice E contém instruções para resolução de exemplos selecionados de cada capítulo usando os programas Maple, Mathematica e MATLAB. A idéia é dar aos estudantes orientação suficiente para que eles possam explorar o poder dos seus sistemas algébricos computacionais, juntamente com as facilidades da ajuda disponível que cada um deles oferece.

Respostas a Exercícios Selecionados de Número Ímpar

Respostas curtas à maioria dos exercícios numéricos de número ímpar são dadas no final do livro.

Agradecimentos

Devo a muitas pessoas que têm, ao longo dos anos, influenciado minha visão sobre álgebra linear e ensino de matemática em geral. Primeiramente, gostaria de agradecer coletivamente aos participantes das sessões especiais de educação e álgebra linear em encontros da Mathematical Association of America e da Canadian Mathematical Society. Também aprendi muito pela participação no Canadian Mathematics Education Study Group.

Quero agradecer especialmente a Ed Barbeau, Bill Higginson, Eric Muller, Morris Orzech, Bill Ralph, Pat Rogers, Peter Taylor e Walter Whiteley por seus conselhos e inspiração e por serem modelos de comportamento marcantes. Agradeço a Jim Stewart primeiramente por me encorajar a escrever este livro e também pelo apoio e sábio aconselhamento ao longo de seu desenvolvimento. Joe Rotman e seu adorável livro *A first course in abstract algebra* (Um primeiro curso em álgebra abstrata) inspiraram as notas etimológicas no livro, e pude contar fortemente com o livro *The words of mathematics* (As palavras da matemática), de Steven Schwartzman, quando redigia essas notas. Edie Gale Hayes e Judith Kelly ajudaram com as citações. Agradecimentos especiais vão para minhas alunas Sarah Jones e Sara Kennedy, por darem assistência com as soluções dos exercícios. Minha colega Wendi Morrison meticulosamente conferiu a precisão e a legibilidade de todo o manuscrito e salvou-me de mais do que uns poucos erros.

Estendo meus agradecimentos a Israel Koltracht, da Connecticut University; Arthur Robinson, da George Washington University; e a Mo Tavakoli, do Chaffey College, que contribuíram com valiosos comentários sobre o texto, do começo à finalização. Estou grato pelo tempo que eles empregaram para fazê-lo. Suas opiniões e visões fizeram deste um livro melhor.

Material com direitos autorais

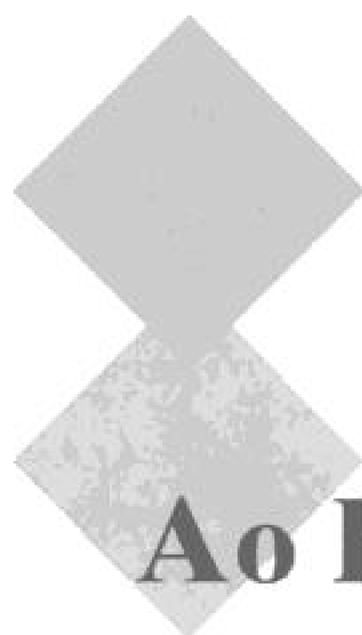
Agradeço aos membros do meu departamento na Trent University por todas as conversas que tivemos sobre matemática; aprendi alguma coisa com cada um deles. Sou grato aos meus alunos por levantarem boas perguntas, proporcionando-me o necessário retorno para tornar-me um professor melhor. Minha maravilhosa secretária, Carolyn Johns, ajudou-me de maneiras incontáveis – enviando fax, e-mails e cartas, pacotes por mensageiros –, mas principalmente por ser ela mesma, eficiente e agradável como de costume.

Agradeço sinceramente a todas as pessoas que estiveram envolvidas na produção deste livro. Agradecimentos especiais vão para Quica Ostrander, da Lifland et. al.; Bookmakers, por tornarem meu manuscrito não refinado em um livro; e George Morris, da Scientific Illustrators, por produzir o tremendo trabalho de arte. Tem sido um prazer trabalhar com toda a produção e a equipe de marketing da Brooks/Cole. Meus agradecimentos vão para Molly Nance, Carol Benedict, Stacy Green, Sue Ewing, Karin Sandberg e Keith Faivre. Tive a sorte de ter trabalhado com dois maravilhosos editores na Brooks/Cole. Gary Ostedt teve a coragem de assinar contrato com um autor inexperiente e a paciência de agüentar meu perfeccionismo crescente enquanto eu tentava escrever o melhor livro possível. Bob Pirtle herdou o projeto e deu-me exatamente o tipo certo de conselho, encorajamento e a pressão gentil necessária para ver o livro sendo completado. Obrigado a ambos.

Finalmente, gostaria de agradecer à minha família, especialmente Marg, por seu amor, paciência e compreensão durante os vários anos que gastei escrevendo este livro em meu "pitoresco escritório subterrâneo".

David Poole
dpoole@trentu.ca

Material com direitos autorais



Ao Professor

"Você poderia dizer-me, por favor, por onde eu
deveria seguir daqui?"
"Isso depende muito de aonde
você pretende chegar", disse o Gato.

Lewis Carroll
Aventuras de Alice no País das Maravilhas, 1865

Este texto foi escrito tendo em mente a flexibilidade. É destinado a um curso de um ou dois semestres, com 36 aulas por semestre. A gama de tópicos e aplicações o torna adequado para uma variedade de público e tipos de curso. No entanto, há mais material no livro do que pode ser coberto em sala de aula, mesmo em um curso de dois semestres. Depois da descrição geral do texto, vêm em seguida algumas sugestões breves de modos de utilizar o livro.

Uma Descrição Geral do Texto

Capítulo 1: Vetores

O jogo da pista de corrida na Seção 1.1 serve como introdução aos vetores de um modo informal. (Também é divertido de jogar!) Os vetores são depois introduzidos formalmente de dois pontos de vista: um algébrico e um geométrico. As operações de adição e multiplicação por escalar e suas propriedades são primeiramente desenvolvidas no caso concreto de $\mathbb{R}^2$ e $\mathbb{R}^3$ antes de serem generalizadas para $\mathbb{R}^n$. A Seção 1.3 define o produto escalar e as noções a ele relacionadas de comprimento, ângulo e ortogonalidade, e desenvolve o conceito muito importante de projeção (ortogonal), que reaparece nos Capítulos 5 e 7. A investigação "Vetores e Geometria" mostra como métodos vetoriais podem ser usados para demonstrar certos resultados de geometria euclidiana. A Seção 1.4 é uma introdução básica, porém meticulosa, a retas e planos em $\mathbb{R}^2$ e $\mathbb{R}^3$; é crucial para a compreensão do significado geométrico das soluções de sistemas lineares no Capítulo 2. Note que o produto vetorial de vetores em $\mathbb{R}^3$ é deixado como uma investigação. O capítulo termina com uma introdução aos códigos, levando a uma discussão sobre aritmética modular e álgebra linear finita. A maioria dos alunos irá gostar da aplicação ao UPC (Universal Product Code) e ao ISBN (International Standard Book Number).

Material com direitos autorais

Capítulo 2: Sistemas de Equações Lineares

A introdução ao Capítulo 2 ilustra que há mais de uma maneira de pensar na solução de um sistema de equações lineares. As Seções 2.2 e 2.3 desenvolvem a principal ferramenta computacional para resolução de sistemas lineares: redução de matrizes por linhas (métodos de eliminação de Gauss e de Gauss-Jordan). Quase todos os métodos computacionais subseqüentes neste livro dependem dessa ferramenta. O Teorema do Posto aparece aqui pela primeira vez; ele aparece novamente, com mais generalidade, nos Capítulos 3, 5 e 6. A Seção 2.4 é muito importante: introduz as noções fundamentais de conjuntos geradores e independência linear de vetores. Não se apresse ao passar por esse tópico. A Seção 2.4 contém cinco aplicações, dentre as quais os professores podem escolher algumas, dependendo do tempo disponível e do interesse da classe. Os métodos iterativos na Seção 2.6 serão opcionais para muitos cursos, mas essenciais para um curso com um foco aplicado/numérico. O que as três investigações deste capítulo têm em comum é que todas elas lidam com aspectos sobre o uso de computadores para a resolução de sistemas lineares. Todos os estudantes devem pelo menos ser conscientizados quanto a essas questões.

Capítulo 3: Matrizes

O Capítulo 3 contém algumas das idéias mais importantes do livro. É um capítulo longo, mas o material inicial pode ser coberto de modo razoavelmente rápido, permitindo um tempo extra para o material crucial da Seção 3.5. A Seção 3.1 é uma investigação que introduz a noção de transformação linear: a idéia de que matrizes não são apenas objetos estáticos, mas um tipo de função, transformando vetores em outros vetores. Todos os fatos básicos sobre matrizes, operações com matrizes e suas propriedades são encontrados nas duas primeiras seções. Vale a pena dar ênfase ao material sobre matrizes decompostas e as múltiplas representações do produto de matrizes, pois ele será usado repetidamente em seções subseqüentes. O Teorema Fundamental das Matrizes Invertíveis, na Seção 3.4, é muito importante e irá aparecer outras vezes, quando novas caracterizações de inversibilidade forem apresentadas. O objetivo da Seção 3.5 é apresentar vários conceitos-chave da álgebra linear (subespaço, base, dimensão e posto) no contexto concreto de matrizes antes que os alunos os vejam em toda a generalidade. Embora os exemplos nessa seção sejam todos familiares, é importante que os alunos se acostumem com a nova terminologia e compreendam o significado de base. Ao se colocar o tratamento geométrico das transformações lineares na Seção 3.6, a intenção é suavizar a transição para transformações lineares em geral, no Capítulo 6. O exemplo de projeção é particularmente importante, pois reaparece no Capítulo 5. Há quatro aplicações na Seção 3.7, entre as quais se pode escolher alguma. Pelo menos as cadeias de Markov ou o modelo de Leslie de crescimento populacional devem ser cobertos de modo que possam ser usados novamente no Capítulo 4, onde seus comportamentos são explicados. Vale a pena propor aos alunos como um projeto, ou apresentar em forma de oficina de estudos, a investigação desse capítulo – sobre a importante fatorização *LU* de uma matriz.

Capítulo 4: Autovalores e Autovetores

Para manter a ênfase geométrica do livro, a Seção 4.2 usa a original apresentação de "autofiguras" como uma forma de ajudar os estudantes a visualizar os autovetores de matrizes 2×2. Determinantes aparecem na Seção 4.3, na qual são usados para encontrar os polinômios característicos de matrizes pequenas. Esse "curso intensivo" sobre determinantes contém todo o material necessário de que os estudantes precisam, incluindo uma opcional, porém elementar, demonstração do Teorema de Expansão de Laplace. A investigação "Aplicações Geométricas dos Determinantes" contém vários resultados interessantes e úteis e produz um projeto interessante. (Professores que queiram dar uma cobertura mais detalhada dos determinantes podem optar por mostrar parte dessa investigação em sala de aula.) A Seção 4.4 apresenta a teoria básica de autovalores e autovetores; a Seção 4.5 trata do importante tópico sobre diagonalização. Vale a pena cobrir em sala de aula o Exemplo 8 da Seção 4.5, sobre potências de matrizes.

Material com direitos autorais

O método das potências e suas variações, discutidos na Seção 4.6, são opcionais, mas todos os alunos devem ser conscientizados da existência de tal método. Um curso aplicado deve cobri-lo em detalhes. Cadeias de Markov e o modelo de crescimento populacional de Leslie reaparecem na Seção 4.7. Embora a demonstração do Teorema de Perron-Frobenius seja opcional, o teorema deve ser pelo menos mencionado, pois explica *por que* devemos esperar um único autovalor positivo com um autovetor correspondente positivo nessas aplicações. As aplicações sobre relações de recorrência e equações diferenciais conectam a álgebra linear à matemática discreta e ao cálculo diferencial e integral, respectivamente. A matriz exponencial pode ser abordada em uma turma com uma boa base de cálculo diferencial e integral.

Capítulo 5: Ortogonalidade

A investigação introdutória "Sombras em uma Parede" é matemática da melhor qualidade: toma um conceito conhecido (projeção de um vetor sobre um outro vetor) e o generaliza de uma maneira útil (projeção de um vetor sobre um subespaço – um plano) enquanto revela algumas propriedades adicionais. A Seção 5.2 contém os resultados básicos sobre conjuntos de vetores ortogonais e ortonormais, que serão usados repetidamente daqui para a frente. Em particular, será dada ênfase às matrizes ortogonais. A Seção 5.3 generaliza dois conceitos do Capítulo 1: o complemento ortogonal de um subespaço e a projeção ortogonal de um vetor sobre um subespaço. O Teorema da Decomposição Ortogonal é importante aqui e ajuda a preparar o Processo de Gram-Schmidt. Note também a demonstração curta do Teorema do Posto. O Processo de Gram-Schmidt é detalhado na Seção 5.4, juntamente com a fatoração QR, que é extremamente importante. Duas investigações delineiam a maneira como a fatoração QR é calculada na prática e como ela pode ser utilizada para achar aproximações de autovalores. A Seção 5.5, sobre diagonalização ortogonal de matrizes (reais) simétricas, é necessária para as aplicações que se seguem. A seção também contém o Teorema Espectral, um dos destaques da álgebra linear. As aplicações na Seção 5.6 incluem códigos duplos, formas quadráticas e esboço de gráficos de equações quadráticas. Eu sempre incluo pelo menos a última dessas aplicações em meu curso, pois ela estende o que os alunos já conhecem sobre seções cônicas.

Capítulo 6: Espaços Vetoriais

A maioria dos estudantes irá apreciar o aprendizado sobre os quadrados mágicos na Seção 6.1. O Problema 4 está relacionado mais diretamente com o tópico desse capítulo, mas os outros resultados são encaminhados nos Exercícios 46-47, na Seção 6.2, e nos Exercícios 42-43, na Seção 6.3. Como todas as idéias principais sobre espaços vetoriais já foram apresentadas nos Capítulos 1 a 3, os estudantes devem achar as Seções 6.2 e 6.3 bastante familiares. A ênfase aqui deve ser o uso dos axiomas de espaço vetorial para demonstrar as propriedades, em vez de confiar em técnicas computacionais. Ao discutir mudança de base, na Seção 6.4, é útil mostrar aos estudantes como usar a notação para lembrar como a construção funciona. Por fim, o método de Gauss-Jordan é o mais eficiente aqui. As Seções 6.5 e 6.6, sobre transformações lineares, são importantes. Os exemplos se relacionam com os resultados prévios sobre matrizes (e matrizes de transformações). Em particular, é importante enfatizar que os conceitos de núcleo e de imagem de uma transformação linear são generalizações dos conceitos de núcleo e de espaço coluna de uma matriz. A Seção 6.7 apresenta a noção de que (quase) toda transformação linear é essencialmente uma matriz de transformação. Como essa idéia foi semeada na Seção 3.6, os estudantes não devem achá-la terrivelmente surpreendente. Entretanto, os exemplos devem ser trabalhados cuidadosamente.

A conexão entre mudança de base e semelhança de matrizes é notável. A investigação "Mosaicos, Reticulados e a Restrição Cristalográfica" é uma aplicação admirável da mudança de base. A conexão com o trabalho artístico de M. C. Escher o torna ainda mais interessante. As aplicações na Seção 6.8 somam-se às anteriores e podem ser incluídas conforme o tempo e o interesse permitirem.

Material com direitos autorais

Capítulo 7: Distância e Aproximação

A Seção 7.1 contém a interessante investigação sobre a "A Geometria do Taxista". Seu objetivo é preparar o material sobre as funções norma e distância (métrica) que aparecem em seguida. Espaços com produto interno são discutidos na Seção 7.2; deve-se dar ênfase aqui (a) aos exemplos e (b) ao uso dos axiomas. Os estudantes geralmente gostam da investigação seguinte, "Desigualdades Geométricas e Problemas de Otimização". (Eles se divertem ao ver como muitos problemas de cálculo diferencial e integral podem ser resolvidos sem usar nada de cálculo!) A Seção 7.3 cobre normas de vetores e de matrizes e mostra como o número condicional de uma matriz se relaciona com a noção de sistemas lineares malcondicionados estudados no Capítulo 2. O método dos mínimos quadrados é uma aplicação importante da álgebra linear em muitas outras disciplinas. O Teorema da Melhor Aproximação e o Teorema dos Mínimos Quadrados são importantes, mas suas demonstrações são intuitivamente claras. Empregue o tempo aqui com alguns exemplos. A decomposição por valores singulares (Seção 7.5) é uma das mais admiráveis aplicações da álgebra linear. Se seu curso conseguir chegar até esse ponto, você será grandemente recompensado. Não apenas a decomposição liga várias noções discutidas previamente; ela também proporciona algumas novas (e muito poderosas) aplicações. Se o programa MATLAB estiver disponível, vale a pena apresentar a aplicação à compressão de imagens digitais – é uma exposição visualmente impressionante do poder da álgebra linear e uma coroação adequada ao curso. Na Seção 7.6, mais aplicações podem ser escolhidas com base no tempo disponível e no interesse da classe.

Como Usar o Livro

Os estudantes acham que este livro é de leitura fácil, por isso geralmente eu os faço ler uma seção antes de cobrir o assunto em sala de aula. Dessa forma, posso gastar mais tempo enfatizando os conceitos mais importantes, lidando com tópicos em que os estudantes têm mais dificuldade, explorando exemplos e discutindo aplicações. Não tento cobrir todo o material da leitura feita em casa durante a aula. Essa aproximação permite-me manter o passo do curso de modo bastante rápido, desacelerando nas seções que, tipicamente, os estudantes acham mais desafiadoras.

Em um curso de dois semestres, é possível cobrir quase o livro inteiro, incluindo uma razoável seleção de aplicações. Para obter uma flexibilidade extra, você pode omitir alguns dos tópicos (por exemplo, dar apenas um breve tratamento de álgebra linear numérica), liberando assim tempo para uma cobertura mais profunda dos demais tópicos, de mais aplicações ou de alguma das investigações. Em cursos que exijam maior profundidade em matemática e que enfatizem as demonstrações, muito do material dos Capítulos 1 a 3 pode ser coberto rapidamente. O Capítulo 6 pode então ser coberto em conjunto com as Seções 3.5 e 3.6, e o Capítulo 7 pode ser integrado ao Capítulo 5. Para tais cursos, eu garantiria que os alunos fizessem as investigações dos Capítulos 1, 4, 6 e 7.

Para um curso de um semestre, a natureza do curso e o público-alvo determinam quais tópicos incluir. Três tipos possíveis de curso são descritos a seguir. Cada um tem menos do que as 36 horas sugeridas, permitindo tempo para tópicos extras, revisões e provas.

Um Curso Básico

Um curso concebido para alunos de graduação em Matemática e estudantes de outros cursos é esboçado a seguir. Esse curso não menciona espaço vetorial geral de nenhuma forma (todos os conceitos são tratados em casos concretos) e é muito leve em demonstrações. Mesmo assim, é uma introdução bastante cuidadosa à álgebra linear.

Material com direitos autorais

Capítulo	Seções	Número de aulas
1	1 a 3	4
2	1 a 3	4
3	1 a 5	8
4	1 a 4	6
5	1, 2, 4	4
7	3	2
		28 aulas

Como o conjunto de estudantes de uma disciplina como essa representa uma variedade grande de cursos, eu sugeriria o uso das aulas remanescentes para aplicações. Em meus cursos, eu cubro a Seção 1.5, de que os alunos realmente parecem gostar, e pelo menos uma aplicação de cada um dos Capítulos 2 a 5. Outras aplicações podem ser atribuídas como projetos, juntamente com tantas investigações quantas forem desejáveis.

Um Curso com uma Componente Numérica

Para um curso com uma componente numérica, o curso básico esboçado há pouco pode ser suplementado com as seções do texto que lidam com álgebra linear numérica. Em tais cursos, eu cobriria parcial ou totalmente as Seções 2.6, 4.6, 5.4, 7.3 e 7.5, terminando com a decomposição por valores singulares. As investigações nos Capítulos 2, 3 e 5 são particularmente convenientes para tais cursos, bem como quase qualquer uma das aplicações.

Um Curso para Estudantes que Já Estudaram Alguma Álgebra Linear

Alguns cursos serão dirigidos a estudantes que já depararam com os princípios da álgebra linear em outros cursos. Por exemplo, uma disciplina de Álgebra Básica irá freqüentemente incluir uma introdução a sistemas de equações lineares, matrizes e determinantes; uma disciplina de cálculo diferencial com várias variáveis quase certamente terá material sobre vetores, retas e planos. Para estudantes que já viram tais tópicos, muito do material do início pode ser omitido e trocado por uma rápida revisão. Dependendo da experiência anterior dos alunos da classe, pode ser possível apresentar rapidamente o material até a Seção 3.4, em cerca de seis aulas. Se a classe tiver um número significativo de alunos de graduação em Matemática (e especialmente se este for o único curso de Álgebra Linear que eles terão), eu iria assegurar que fossem cobertas as Seções 1.5, 6.2 a 6.6, 7.5 e tantas aplicações quanto o tempo permitisse. Se a classe tiver estudantes de graduação da área de Ciências Exatas (mas não em Matemática), eu iria cobrir as Seções 1.5, 6.2, 7.2 e uma larga seleção de aplicações, tendo a certeza de incluir o material sobre equações diferenciais e aproximações de funções. Se alunos de Ciência da Computação ou Engenharia forem proeminentemente representados, eu tentaria apresentar o máximo possível do material sobre códigos e álgebra linear numérica.

Há muitos outros tipos de curso que podem usar este texto com sucesso. Espero que você o ache útil para o seu curso e que goste de usá-lo.

Material com direitos autorais

Ao Estudante

"Por favor, sua Majestade, por onde devo
começar?", ele perguntou.
"Comece pelo começo", disse o rei, gravemente,
"e continue até que você chegue ao final:
então pare."

Lewis Carroll
Aventuras de Alice no País das Maravilhas, 1865

Álgebra linear é um assunto empolgante. Ele tem muitos resultados interessantes, aplicações em outras disciplinas e conexões com outras áreas da matemática. Veja a seguir algumas sugestões gerais para uso deste livro.

A álgebra linear tem muitas faces: há as *técnicas computacionais*, os *conceitos* e as *aplicações*. Uma das metas deste livro é ajudá-lo a dominar todas essas facetas do assunto e perceber a relação entre elas. Conseqüentemente, é importante que você leia e entenda cada seção do texto antes de tentar fazer os exercícios. Se você ler apenas os exemplos que estão relacionados com os exercícios atribuídos como lição de casa, poderá não entender muito. Certifique-se de que você compreende as definições dos termos e o significado dos teoremas. Não se preocupe se tiver que ler alguma coisa mais de uma vez antes de compreendê-la. Enquanto isso, tenha à mão um lápis e uma calculadora. Pare para fazer você mesmo os exemplos ou para completar os cálculos que estiverem incompletos. O símbolo ➡ na margem indica um lugar onde você deve fazer uma pausa e pensar novamente sobre o que leu até aquele momento.

As respostas à maioria dos exercícios numéricos de número ímpar estão no final do livro. Resista à tentação de olhar a resposta antes de completar um exercício. Lembre-se de que, mesmo quando sua resposta for diferente da apresentada no fim do livro, ela ainda poderá estar correta; há mais de uma maneira correta de expressar algumas das soluções. Por exemplo, um valor de $1/\sqrt{2}$ pode também ser expresso como $\sqrt{2}/2$, e o conjunto de todos os múltiplos escalares do vetor $\begin{bmatrix} 3 \\ 1/2 \end{bmatrix}$ é o mesmo que o conjunto de todos os múltiplos escalares de $\begin{bmatrix} 6 \\ 1 \end{bmatrix}$.

Quando encontrar um conceito novo, tente relacioná-lo com exemplos que você conhece. Escreva as demonstrações e as soluções dos exercícios de uma maneira lógica, conexa, usando sentenças completas. Leia novamente o que você escreveu para ver se faz sentido. Melhor ainda: se puder, peça a um amigo de sua classe para ler o que você escreveu. Se não fizer sentido para outro leitor, muito provavelmente não faz sentido mesmo.

Você irá achar útil uma calculadora com capacidade de operar com matrizes ou um sistema algébrico computacional. Essas ferramentas podem ajudá-lo a conferir suas próprias contas feitas à mão e são indispensáveis para problemas que envolvem contas tediosas. A tecnologia também permitirá que você explore certos aspectos da álgebra linear por você mesmo. Você pode jogar "O que aconteceria se...?": o que aconte-

Material com direitos autorais

ceria se eu mudasse uma das coordenadas neste vetor? O que aconteceria se esta matriz fosse de um tamanho diferente? Posso forçar para que a solução seja como eu gostaria, mudando alguma coisa?

Para sinalizar lugares no texto ou exercícios nos quais o uso de tecnologia é recomendado, coloquei o sinal CAS na margem. O Apêndice E contém exemplos selecionados do livro calculados usando Maple, Mathematica e MATLAB.

Você está prestes a embarcar em uma jornada por meio da álgebra linear. Pense neste livro como um guia de viagem. Você está pronto? Vamos!

Material com direitos autorais



1 Vetores

Lá vêm elas como se viessem do nada.
Pequenas flechas para mim e para você.

– Albert Hammond e Mike Hazelwood
Little arrows
Dutchess Music/BMI, 1968

1.1 Introdução: O Jogo da Pista de Corrida

Muitas quantidades mensuráveis – tais como comprimento, área, volume, massa e temperatura – podem ser completamente descritas pela especificação de sua magnitude. Outras quantidades – como velocidade, força e aceleração – requerem tanto uma magnitude como uma direção para serem descritas. Essas quantidades são *vetores*. Por exemplo, a velocidade do vento é um vetor que consiste na intensidade do vento e da sua direção, tal como 10 km/h na direção sudoeste. Geometricamente, os vetores são freqüentemente representados por flechas ou segmentos de reta orientados.

Embora a idéia de vetor tenha sido introduzida no século XIX, sua utilidade em aplicações – particularmente as aplicações em ciências físicas – não foi percebida até o século XX. Mais recentemente, vetores tiveram aplicações em ciência da computação, estatística, economia e ciências sociais. Examinaremos algumas dessas muitas aplicações ao longo deste livro.

Este capítulo introduz a noção de vetores e começa a examinar algumas de suas propriedades geométricas e algébricas. Examinaremos também uma aplicação não geométrica na qual vetores são úteis. Começaremos, entretanto, com um jogo simples que introduz algumas das idéias cruciais. [Você pode até querer jogá-lo com um amigo durante aqueles (muito raros!) momentos chatos da aula de Álgebra Linear.]

O jogo acontece em um papel quadriculado. Uma pista, com uma linha de partida e uma linha de chegada, é desenhada no papel. A pista pode ser de qualquer comprimento e forma, desde que seja suficientemente larga para acomodar todos os jogadores. Neste exemplo, temos dois jogadores (vamos chamá-los de Ana e Beto) que usam canetas de cores diferentes para representar seus carros ou bicicletas, ou outra coisa que eles usem para percorrer a pista. (Vamos pensar em Ana e Beto como ciclistas.)

Ana e Beto começam desenhando uma marca sobre a linha de partida, em um dos pontos da grade do papel quadriculado. Eles se revezam para avançar para um novo ponto da grade, de acordo com as seguintes regras:

1. Cada novo ponto da grade e o segmento de reta que o liga ao ponto anterior precisam estar inteiramente dentro da pista.

Material com direitos autorais

2. Dois jogadores não podem ocupar o mesmo ponto da grade ao mesmo tempo. (Esta é a regra que proíbe colisões.)
3. Cada novo movimento está relacionado com o movimento anterior da seguinte maneira: se em um movimento um jogador anda a unidades horizontalmente e b unidades verticalmente, então, em seu próximo movimento, esse jogador deve andar entre $a - 1$ e $a + 1$ unidades horizontalmente, e entre $b - 1$ e $b + 1$ unidades verticalmente. Em outras palavras, se o segundo movimento é de c unidades horizontalmente e d unidades verticalmente, então $|a - c| \leq 1$ e $|b - d| \leq 1$. (Esta é a regra da "aceleração/desaceleração".) Note que esta regra obriga o primeiro movimento a ser de uma unidade verticalmente e/ou de uma unidade horizontalmente.

É eliminado o jogador que colide com outro ou sai da pista. O vencedor é o primeiro jogador que cruza a linha de chegada. Se mais de um jogador cruzar a linha de chegada na mesma vez, aquele que ultrapassar mais a linha de chegada será o vencedor.

> O matemático irlandês William Rowan Hamilton (1805–1865) usou conceitos de vetores em seu estudo dos números complexos e das suas generalizações, os quatérnios.

No exemplo de jogo mostrado na Figura 1, a vencedora foi Ana. Beto acelerou demais e teve dificuldade para fazer a curva na parte superior da pista.

Para entender a regra 3, considere o terceiro e o quarto movimentos de Ana. Em seu terceiro movimento, ela andou uma unidade horizontalmente e três unidades verticalmente. Em seu quarto movimento, as opções que ela tinha eram andar de zero a duas unidades horizontalmente e de duas a quatro unidades verticalmente. (Note que algumas dessas combinações a teriam levado para fora da pista.) Ela escolheu andar duas unidades em cada direção.

Problema 1: Jogue alguns jogos de pista de corrida.

Problema 2: Seria possível Beto vencer essa corrida escolhendo uma seqüência diferente de movimentos?

Problema 3: Qual é o menor número de movimentos necessário para retornar a um ponto no qual você já esteve antes?

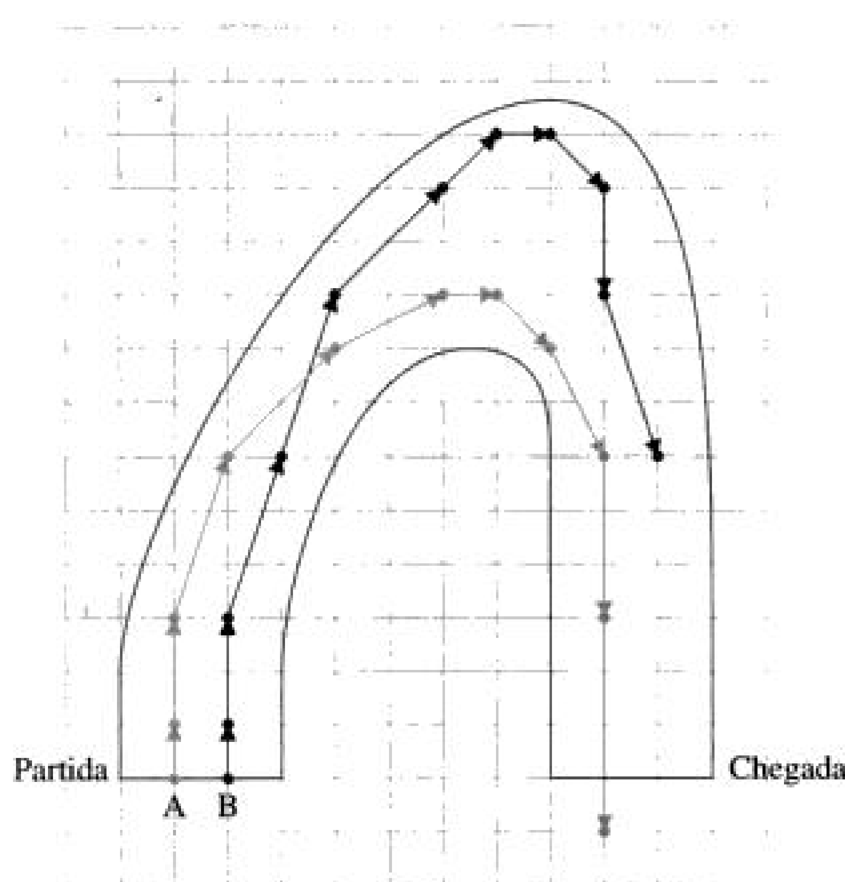


Figura 1 Um exemplo de jogo de pista de corrida

Material com direitos autorais

Problema 4: Use a notação $[a, b]$ para denotar um movimento de a unidades horizontalmente e b unidades verticalmente. (Tanto a como b podem ser negativos.) Se o movimento $[3, 4]$ acabou de ser feito, desenhe no papel quadriculado todos os pontos da grade que podem ser alcançados no próximo movimento.

Problema 5: Qual o efeito *líquido* de dois movimentos sucessivos? Em outras palavras, se você andar $[a, b]$ e depois $[c, d]$, quanto você andará horizontalmente e verticalmente no total?

Problema 6: Escreva a seqüência de movimentos de Ana usando a notação $[a, b]$. Suponha que ela comece na origem $(0, 0)$ dos eixos de coordenadas. Explique como você pode achar as coordenadas dos pontos da grade correspondentes a cada um dos movimentos dela, *sem olhar para o papel quadriculado*. Se os eixos tivessem sido traçados de um outro jeito, de modo que o ponto de partida de Ana fosse o ponto $(2, 3)$, e não a origem, quais seriam as coordenadas do ponto onde ela parou?

Embora simples, esse jogo introduz várias idéias que serão úteis em nosso estudo de vetores. As três próximas seções consideram vetores dos pontos de vista geométrico e algébrico, começando como no jogo de pista de corrida, no plano.

1.2 A Geometria e a Álgebra dos Vetores

Vetores no Plano

Começamos considerando o plano Cartesiano com os conhecidos eixos x e y.[1] Um ***vetor*** é um *segmento de linha orientado* que corresponde ao *deslocamento* de um ponto A até outro ponto B, conforme mostra a Figura 1 a seguir.

A palavra *vetor* vem de um radical latino que significa "carregar". Um vetor é formado quando um ponto é deslocado – ou "carregado" – por uma certa distância em uma certa direção. Visto de outro modo, um vetor "carrega" duas peças de informação: seu comprimento e sua direção.

O vetor de A até B é denotado por $\overrightarrow{AB}$; dizemos que o ponto A é o ***ponto inicial*** ou ***origem*** desse vetor, e que o ponto B é o seu ***ponto final*** ou ***extremidade***. Muitas vezes, um vetor é simplesmente denotado por uma só letra minúscula em negrito, como $\mathbf{v}$.[2]

O conjunto de todos os pontos do plano corresponde ao conjunto de todos os vetores cujos pontos iniciais coincidem com a origem O do plano Cartesiano. A cada ponto A corresponde um vetor $\mathbf{a} = \overrightarrow{OA}$; a cada vetor $\mathbf{a}$ com ponto inicial em O corresponde seu ponto final A. (Vetores com essa forma são às vezes chamados *vetores posição*.)

É natural representar tais vetores usando coordenadas. Por exemplo, na Figura 2, $A = (3, 2)$, e escrevemos o vetor $\mathbf{a} = \overrightarrow{OA} = [3, 2]$ usando colchetes. De modo análogo, os outros vetores da Figura 2 são

$$\mathbf{b} = [-1, 3] \text{ e } \mathbf{c} = [2, -1]$$

[1] O plano cartesiano recebeu esse nome em homenagem ao filósofo matemático francês René Descartes (1596-1650), que introduziu o conceito de coordenadas. Esse conceito permitiu que problemas *geométricos* fossem tratados com o uso de técnicas *algébricas*.

[2] É difícil indicar negrito quando escrevemos vetores à mão. Algumas pessoas preferem escrever $\vec{v}$ para representar o vetor denotado por $\mathbf{v}$ no texto impresso, mas na maioria dos casos é aceitável usar simplesmente a letra v minúscula. Geralmente ficará claro, pelo contexto, se essa letra denota um vetor ou não.

Material com direitos autorais

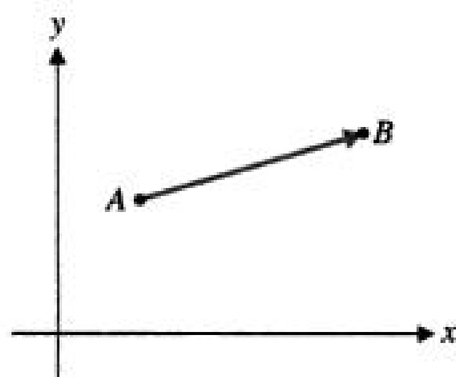


Figura 1

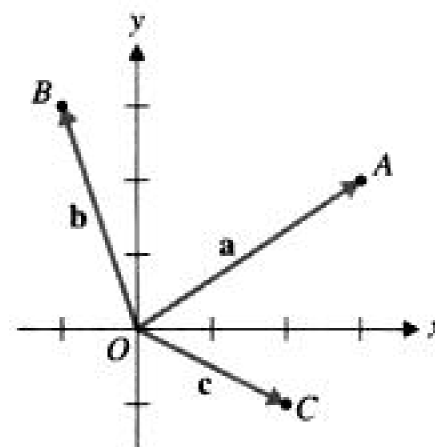


Figura 2

A palavra *componente* é derivada das palavras latinas *co*, que significa "junto", e *ponere*, que significa "pôr". Um vetor é, portanto, formado por suas componentes postas uma junto da outra.

As coordenadas individuais (3 e 2, no caso de **a**) são chamadas ***componentes**** do vetor. Às vezes nos referimos a um vetor como um *par ordenado* de números reais. A ordem é importante, pois, por exemplo, $[3, 2] \neq [2, 3]$. Em geral, dois vetores são iguais se e somente se suas componentes correspondentes forem iguais. Assim, $[x, y] = [1, 5]$ implica que $x = 1$ e $y = 5$.

Muitas vezes é conveniente usar ***vetores coluna*** em vez de (ou além de) ***vetores linha***. Outra representação de $[3, 2]$ é $\begin{bmatrix} 3 \\ 2 \end{bmatrix}$. (O ponto importante é que as componentes estão *ordenadas*.) Em capítulos futuros, você verá que vetores coluna são um pouco melhores do ponto de vista computacional. Por ora, tente se acostumar com ambas as representações.

Você talvez tenha notado que não podemos realmente desenhar o vetor $[0, 0] = \overrightarrow{OO}$ da origem até ela mesma. Mesmo assim, esse é um vetor tão bom quanto qualquer outro e tem um nome especial: ***vetor nulo***. O vetor nulo é denotado por $\mathbf{0}$.

O conjunto de todos os vetores com duas componentes é denotado por $\mathbb{R}^2$ ($\mathbb{R}$ denota o conjunto de todos os números reais, dentre os quais as coordenadas dos vetores em $\mathbb{R}^2$ são escolhidas).[3] Assim, $[-1, 3{,}5]$, $[\sqrt{2}, \pi]$ e $\mathbf{c} = [\frac{5}{3}, 4]$ estão todos em $\mathbb{R}^2$.

Pensando novamente no jogo da pista de corrida, tentemos relacionar aquelas idéias com vetores cujos pontos iniciais *não* estão na origem. O fato de a palavra *vetor* ter sua raiz etimológica no verbo "carregar" nos dá uma pista. O vetor $[3, 2]$ pode ser interpretado da seguinte maneira: começando na origem O, ande três unidades para a direita e depois duas unidades para cima, parando em P. O mesmo deslocamento pode ser aplicado com outros pontos iniciais. A Figura 3 mostra dois deslocamentos equivalentes, representados pelos vetores $\overrightarrow{AB}$ e $\overrightarrow{CD}$.

Dizemos que dois vetores são *iguais* se eles têm o mesmo comprimento, a mesma direção e o mesmo sentido. Assim, $\overrightarrow{AB} = \overrightarrow{CD}$ na Figura 3. (Embora tenham diferentes pontos iniciais e finais, eles representam o mesmo deslocamento.) Geometricamente, dois vetores são iguais se obtemos um deles deslocando o outro paralelamente a si próprio (ou seja, fazendo uma *translação*), até que os dois vetores coincidam. Em termos das componentes, na Figura 3 temos $A = (3, 1)$ e $B = (6, 3)$.[4] Note que o vetor $[3, 2]$, que representa o deslocamento, é simplesmente a diferença das respectivas componentes:

$$\overrightarrow{AB} = [3, 2] = [6 - 3, 3 - 1]$$

* N.T.: A expressão "coordenadas do vetor" também é usada.

[3] $\mathbb{R}^2$ é pronunciado como "r dois".

[4] Quando os vetores são expressos em termos das suas coordenadas, eles estão sendo considerados *analiticamente*.

Material com direitos autorais

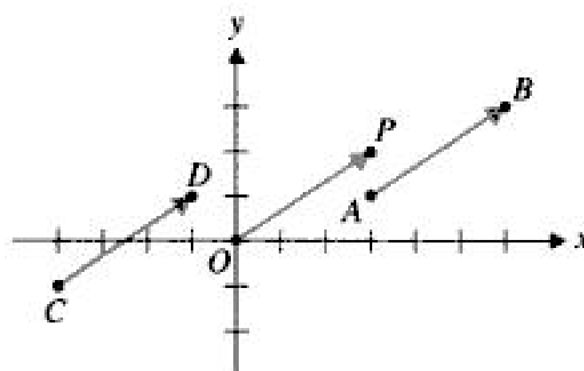


Figura 3

De modo análogo,

$$\overrightarrow{CD} = [-1 - (-4), 1 - (-1)] = [3, 2]$$

e, portanto, $\overrightarrow{AB} = \overrightarrow{CD}$, como se esperava.

Um vetor como $\overrightarrow{OP}$, que tem seu ponto inicial na origem do plano, está na ***posição padrão***. A discussão anterior mostra que todo vetor pode ser desenhado como um vetor na posição padrão. Reciprocamente, um vetor na posição padrão pode ser redesenhado (fazendo-se uma translação) de modo que seu ponto inicial seja um ponto qualquer do plano.

EXEMPLO 1 Dados $A = (-1, 2)$ e $B = (3, 4)$, ache $\overrightarrow{AB}$ e redesenhe-o (a) na posição padrão e (b) com o ponto inicial no ponto $C = (2, -1)$.

SOLUÇÃO: Calculamos $\overrightarrow{AB} = [3 - (-1), 4 - 2)] = [4, 2]$. Se fizermos a translação de $\overrightarrow{AB}$ para $\overrightarrow{CD}$, onde $C = (2, -1)$, precisaremos ter $D = (2 + 4, -1 + 2) = (6, 1)$. (Veja a Figura 4.)

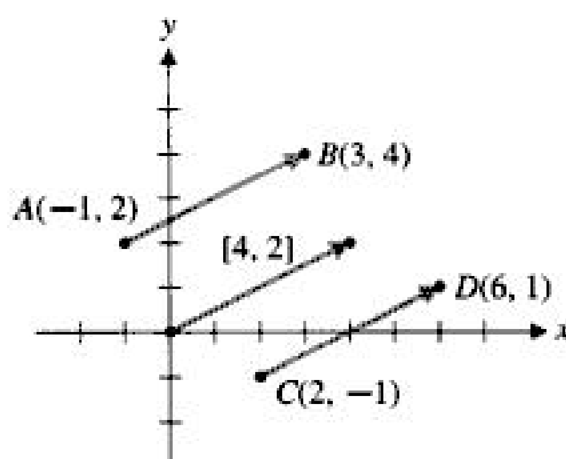


Figura 4

Material com direitos autorais

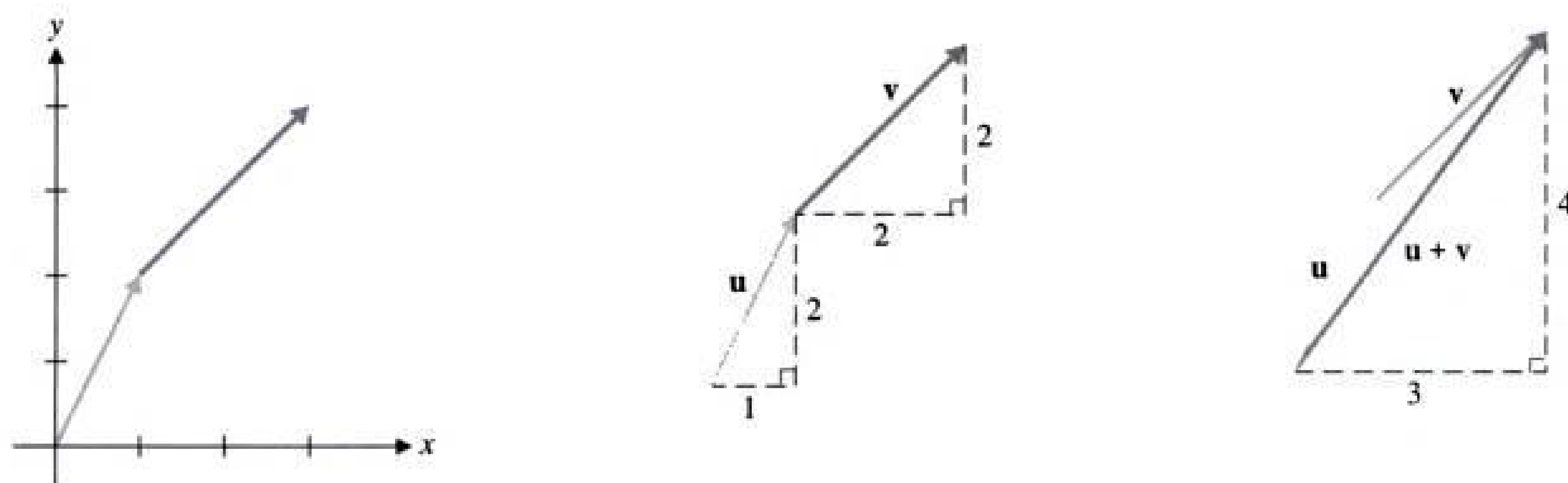


Figura 5 Adição de vetores

Novos Vetores a Partir de Vetores Existentes

Muitas vezes queremos colocar "um vetor depois do outro" e assim fazer um deslocamento suceder outro, como no jogo da pista de corrida. Isso nos leva à noção de ***adição de vetores***, a primeira das operações básicas com vetores.

Se colocarmos **v** depois de **u**, poderemos visualizar o deslocamento total como um terceiro vetor, denotado por **u** + **v**. Na Figura 5, **u** = [1, 2] e **v** = [2, 2], de modo que o efeito resultante de colocar **v** depois de **u** é

$$[1 + 2, 2 + 2] = [3, 4]$$

que representa **u** + **v**. Em geral, se **u** = $[u_1, u_2]$ e **v** = $[v_1, v_2]$, então sua ***soma*** **u** + **v** é o vetor

$$\mathbf{u} + \mathbf{v} = [u_1 + v_1, u_2 + v_2]$$

É útil visualizar **u** + **v** geometricamente. A regra seguinte é a versão geométrica da discussão anterior.

◆ A Regra da Adição

Dados os vetores **u** e **v** no $\mathbb{R}^2$, faça uma translação de **v** de modo que seu ponto inicial coincida com o ponto final de **u**. A ***soma*** **u** + **v** de **u** e **v** é o vetor que sai do ponto inicial de **u** e vai até o ponto final de **v**. (Veja a Figura 6.)

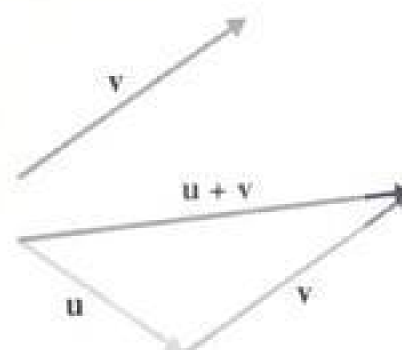


Figura 6 A regra da adição

Material com direitos autorais

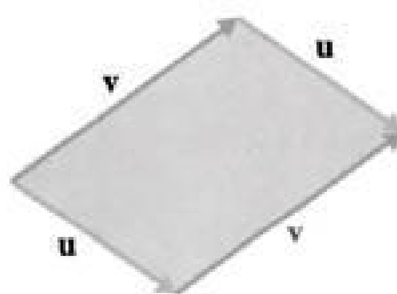


Figura 7 O paralelogramo determinado por **u** e **v**

Transladando **u** e **v** paralelamente a eles mesmos, obtemos um paralelogramo, conforme mostra a Figura 7. Esse paralelogramo é chamado de *paralelogramo determinado por* **u** *e* **v**. Ele nos leva a uma versão equivalente à regra da adição, para vetores na posição padrão.

◆ A Regra do Paralelogramo

Dados os vetores **u** e **v** no $\mathbb{R}^2$ (na posição padrão), sua ***soma*** **u** + **v** é o vetor na posição padrão sobre a diagonal do paralelogramo determinado por **u** e **v**. (Veja a Figura 8.)

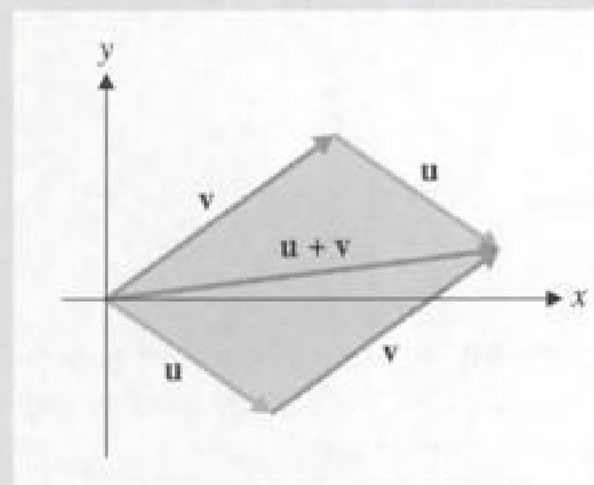
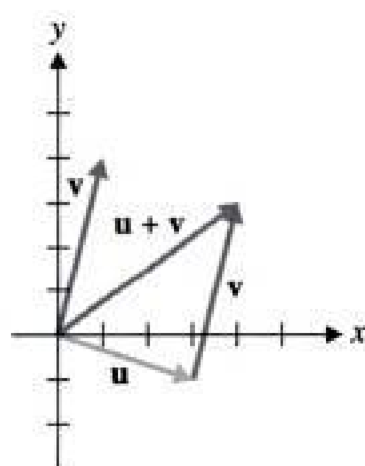


Figura 8 A regra do paralelogramo

EXEMPLO 2 Dados **u** = [3, −1] e **v** = [1, 4], calcule e desenhe **u** + **v**.

SOLUÇÃO: Calculamos **u** + **v** = [3 + 1, − 1 + 4] = [4, 3]. Esse vetor é desenhado por meio da regra da adição na Figura 9(a) e da regra do paralelogramo na Figura 9(b).



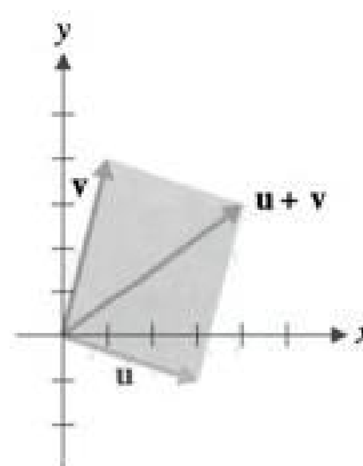


Figura 9 (a) (b)

A segunda operação básica que fazemos com vetores é a ***multiplicação por escalar***. Dado um vetor **v** e um número real c, o ***múltiplo escalar*** c**v** é o vetor obtido pela multiplicação de cada componente de **v** por c. Por exemplo, 3 [−2, 4] = [−6, 12]. Em geral,

Material com direitos autorais

$$c\mathbf{v} = c[v_1, v_2] = [cv_1, cv_2]$$

Geometricamente, $c\mathbf{v}$ é uma versão de **v** em outra escala.

EXEMPLO 3 Se $\mathbf{v} = [-2, 4]$, determine e desenhe $2\mathbf{v}$, $\frac{1}{2}\mathbf{v}$ e $-2\mathbf{v}$.

SOLUÇÃO: Calculamos da seguinte maneira:

$$2\mathbf{v} = [2(-2), 2(4)] = [-4, 8]$$
$$\tfrac{1}{2}\mathbf{v} = [\tfrac{1}{2}(-2), \tfrac{1}{2}(4)] = [-1, 2]$$
$$-2\mathbf{v} = [-2(-2), -2(4)] = [4, -8]$$

Esses vetores são mostrados na Figura 10.

O termo *escalar* vem da palavra grega *scala*, que significa "escada". Os degraus igualmente espaçados de uma escada sugerem uma escala, e, na aritmética vetorial, a multiplicação por uma constante altera apenas a escala (ou comprimento e sentido) de um vetor. Assim, as constantes são conhecidas como escalares.

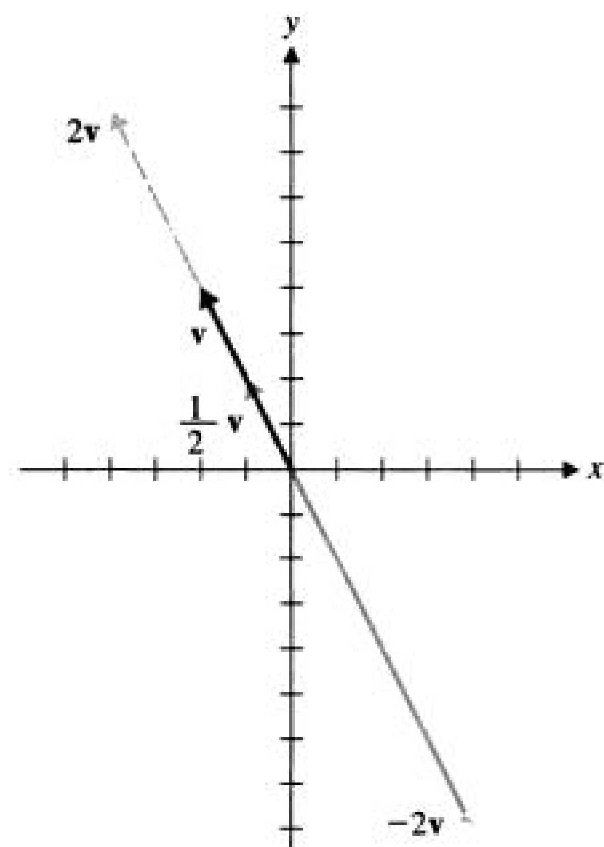


Figura 10 ◆

Observe que $c\mathbf{v}$ tem o mesmo sentido de **v** se $c > 0$, e sentido oposto se $c < 0$. Além disso, o comprimento de $c\mathbf{v}$ é $|c|$ vezes o comprimento de **v**. Por esse motivo, no contexto dos vetores, as constantes (isto é, os números reais) são chamadas de ***escalares***. Levando-se em conta a translação de vetores, dois vetores são múltiplos escalares um do outro se e somente se eles são ***paralelos***, como mostra a Figura 11.

Um caso especial de múltiplo escalar é $(-1)\mathbf{v}$, escrito como $-\mathbf{v}$ e chamado ***oposto de* v**. Podemos usá-lo para definir ***subtração de vetores***: a ***diferença*** entre **u** e **v** é o vetor $\mathbf{u} - \mathbf{v}$, definido por

$$\mathbf{u} - \mathbf{v} = \mathbf{u} + (-\mathbf{v})$$

A Figura 12 mostra que $\mathbf{u} - \mathbf{v}$ corresponde à "outra" diagonal do paralelogramo determinado por **u** e **v**.

Material com direitos autorais

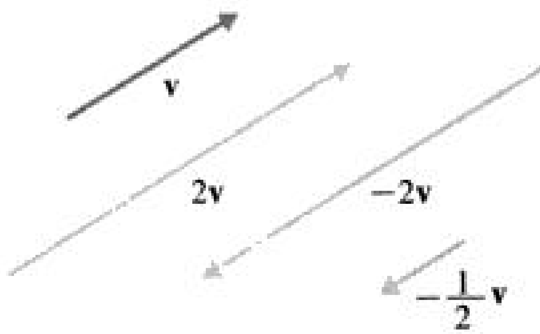
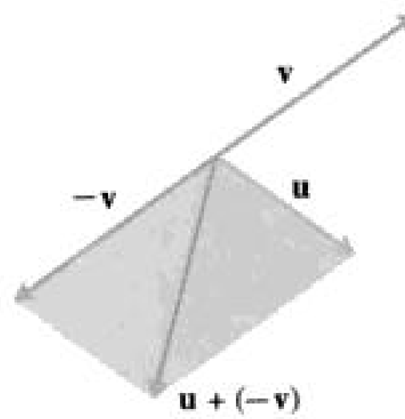


Figura 11



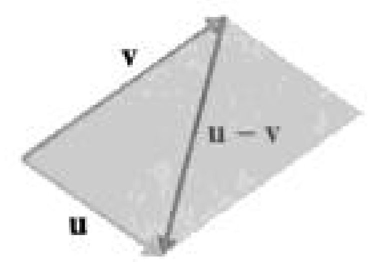
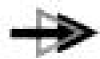


Figura 12 Subtração de vetores

EXEMPLO 4 Se $\mathbf{u} = [1, 2]$ e $\mathbf{v} = [-3, 1]$, então $\mathbf{u} - \mathbf{v} = [1 - (-3), 2 - 1] = [4, 1]$. ◆



A definição de subtração no Exemplo 4 também lembra o procedimento para determinarmos um vetor como $\overrightarrow{AB}$. Se os pontos A e B correspondem aos vetores **a** e **b** na posição padrão, então $\overrightarrow{AB} = \mathbf{b} - \mathbf{a}$, como mostra a Figura 13. [Observe que a regra da adição, aplicada a esse diagrama, fornece a equação $\mathbf{a} + (\mathbf{b} - \mathbf{a}) = \mathbf{b}$. Se tivéssemos, acidentalmente, desenhado $\mathbf{b} - \mathbf{a}$ com extremidade em A e não em B, o diagrama seria lido como $\mathbf{b} + (\mathbf{b} - \mathbf{a}) = \mathbf{a}$, o que é claramente falso! Ainda nesta seção, falaremos mais sobre expressões algébricas que envolvem vetores.]

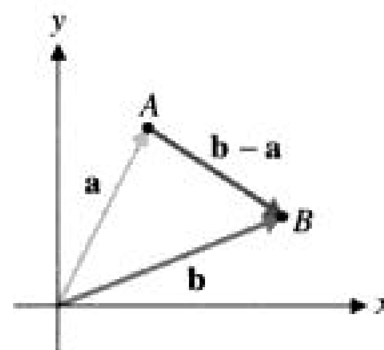


Figura 13

Vetores em $\mathbb{R}^3$

Tudo o que fizemos até agora se generaliza facilmente para três dimensões. O conjunto de todas as *triplas ordenadas* de números reais é denotado por $\mathbb{R}^3$. Pontos e vetores são localizados usando-se três eixos coordenados perpendiculares dois a dois e que se encontram em uma origem O. Por exemplo, o ponto $A = (1, 2, 3)$ pode ser localizado assim: primeiro percorremos uma unidade ao longo do eixo x, então nos movemos duas unidades paralelamente ao eixo y e finalmente percorremos três unidades paralelamente ao eixo z. O vetor correspondente $\mathbf{a} = [1, 2, 3]$ é então $\overrightarrow{OA}$, como mostra a Figura 14.

Uma outra maneira de visualizar o vetor **a** em $\mathbb{R}^3$ é construir uma caixa cujas seis faces sejam determinadas pelos três planos coordenados (os planos xy, xz, e yz) e pelos três planos paralelos aos planos coordenados e que passam por (1, 2, 3). O vetor [1, 2, 3] corresponde à diagonal que liga a origem ao vértice diametralmente oposto (Figura 15).

As definições "componente a componente" da adição de vetores e multiplicação por escalar se estendem a $\mathbb{R}^3$ de maneira óbvia.

Material com direitos autorais

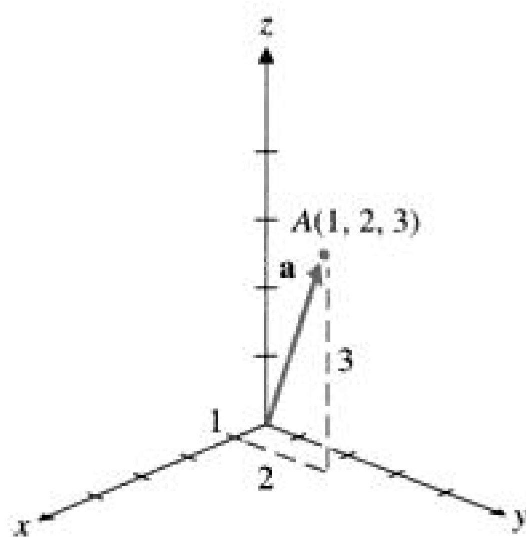


Figura 14

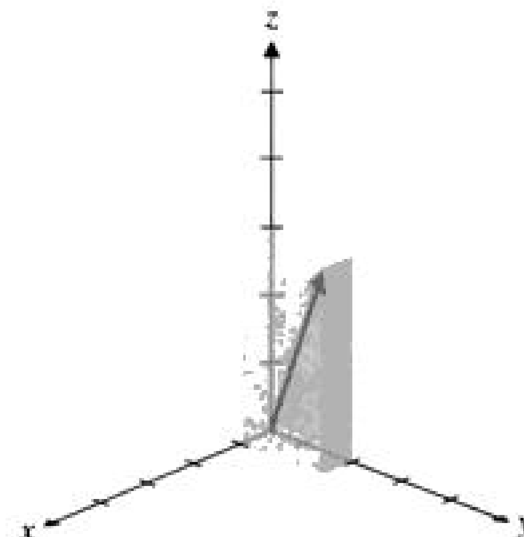


Figura 15

Vetores em $\mathbb{R}^n$

Em geral, definimos $\mathbb{R}^n$ como o conjunto de todas as *n-uplas ordenadas* de números reais, escritos como vetores linha ou coluna. Assim, um vetor **v** em $\mathbb{R}^n$ é da forma

$$[v_1, v_2, \ldots, v_n] \quad \text{ou} \quad \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix}$$

As coordenadas individuais de **v** são as suas componentes; v_i é chamada a i-ésima componente.

Estendemos a definição de adição de vetores e multiplicação por escalar em $\mathbb{R}^n$ de maneira óbvia: se $\mathbf{u} = [u_1, u_2, \ldots, u_n]$ e $\mathbf{v} = [v_1, v_2, \ldots, v_n]$, a i-ésima componente de $\mathbf{u} + \mathbf{v}$ é $u_i + v_i$, e a i-ésima componente de $c\mathbf{v}$ é simplesmente cv_i.

Como não podemos desenhar vetores em $\mathbb{R}^n$, é importante que sejamos capazes de calcular com vetores. Devemos ser cuidadosos e não assumir que a aritmética vetorial seja similar à aritmética dos números reais. Freqüentemente o é, e os cálculos algébricos que fazemos com os vetores são similares aos que fazemos com números. No entanto, nas seções subseqüentes, encontraremos situações em que a álgebra vetorial é *bem diferente* das situações que encontramos em nossa experiência com os números reais. É importante, pois, verificar todas as propriedades algébricas antes de ceder à tentação de usá-las.

Uma dessas propriedades é a *comutatividade* da adição: $\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}$ para vetores **u** e **v**. Isso é certamente verdadeiro em $\mathbb{R}^2$. Geometricamente, a regra de origem-de-um-na-extremidade-do-outro mostra que $\mathbf{u} + \mathbf{v}$ e $\mathbf{v} + \mathbf{u}$ são ambos a mesma diagonal do paralelogramo determinado por **u** e **v**. (A regra do paralelogramo reflete essa simetria; veja a Figura 16.)

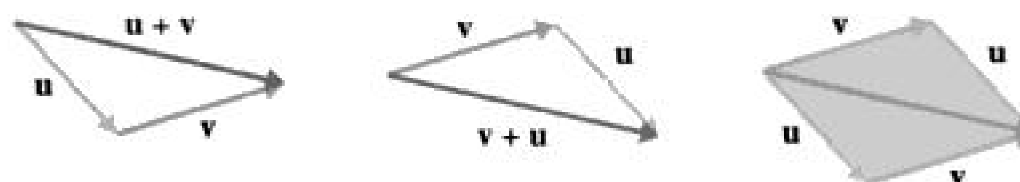
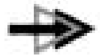


Figura 16 $\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}$

Material com direitos autorais



Note que a Figura 16 é simplesmente uma ilustração da propriedade $\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}$. Não é uma demonstração, pois não considera todos os casos possíveis. Por exemplo, devemos incluir os casos onde $\mathbf{u} = \mathbf{v}$, $\mathbf{u} = -\mathbf{v}$ e $\mathbf{u} = \mathbf{0}$. (Como seria o diagrama para estes casos?) Por esse motivo, é necessária uma demonstração algébrica. Entretanto, é tão fácil fornecer uma demonstração válida em $\mathbb{R}^n$ quanto uma válida em $\mathbb{R}^2$.

O seguinte teorema resume as propriedades algébricas da adição de vetores e multiplicação por escalar em $\mathbb{R}^n$. As demonstrações são conseqüências das propriedades correspondentes dos números reais.

◆ **TEOREMA 1** Propriedades Algébricas dos Vetores em $\mathbb{R}^n$

Sejam $\mathbf{u}$, $\mathbf{v}$ e $\mathbf{w}$ vetores em $\mathbb{R}^n$, e a e b escalares. Então:

a. $\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}$ — Comutatividade
b. $(\mathbf{u} + \mathbf{v}) + \mathbf{w} = \mathbf{u} + (\mathbf{v} + \mathbf{w})$ — Associatividade
c. $\mathbf{u} + \mathbf{0} = \mathbf{u}$
d. $\mathbf{u} + (-\mathbf{u}) = \mathbf{0}$
e. $a(\mathbf{u} + \mathbf{v}) = a\mathbf{u} + a\mathbf{v}$ — Distributividade
f. $(a + b)\mathbf{u} = a\mathbf{u} + b\mathbf{u}$ — Distributividade
g. $a(b\mathbf{u}) = (ab)\mathbf{u}$
h. $1\mathbf{u} = \mathbf{u}$

Observações:

- As propriedades (c) e (d), juntamente com a propriedade (a) de comutatividade, implicam que $\mathbf{0} + \mathbf{u} = \mathbf{u}$, assim como $-\mathbf{u} + \mathbf{u} = \mathbf{0}$.
- Se lermos da direita para a esquerda as propriedades (e) e (f) de distributividade, elas dizem que podemos *fatorar* tanto um escalar quanto um vetor que seja comum às parcelas de uma soma.

A palavra *teorema* se origina da palavra grega *theorema*, que, por sua vez, vem de uma palavra que significa "examinar". Um teorema é baseado em idéias que temos quando observamos exemplos e extraímos deles as propriedades que tentamos provar que valem em geral.

DEMONSTRAÇÃO: Provaremos as propriedades (a) e (b) e deixaremos as demonstrações das outras propriedades como exercícios. Sejam $\mathbf{u} = [u_1, u_2, \ldots, u_n]$, $\mathbf{v} = [v_1, v_2, \ldots, v_n]$ e $\mathbf{w} = [w_1, w_2, \ldots, w_n]$.

(a)
$$\begin{aligned}\mathbf{u} + \mathbf{v} &= [u_1, u_2, \ldots, u_n] + [v_1, v_2, \ldots, v_n] \\ &= [u_1 + v_1, u_2 + v_2, \ldots, u_n + v_n] \\ &= [v_1 + u_1, v_2 + u_2, \ldots, v_n + u_n] \\ &= [v_1, v_2, \ldots, v_n] + [u_1, u_2, \ldots, u_n] \\ &= \mathbf{v} + \mathbf{u}\end{aligned}$$

A segunda e a quarta igualdades valem pela definição de adição de vetores, e a terceira decorre da comutatividade da adição de números reais.

(b) A Figura 17 ilustra a associatividade em $\mathbb{R}^2$. Algebricamente, temos:

$$\begin{aligned}(\mathbf{u} + \mathbf{v}) + \mathbf{w} &= ([u_1, u_2, \ldots, u_n] + [v_1, v_2, \ldots, v_n]) + [w_1, w_2, \ldots, w_n] \\ &= [u_1 + v_1, u_2 + v_2, \ldots, u_n + v_n] + [w_1, w_2, \ldots, w_n] \\ &= [(u_1 + v_1) + w_1, (u_2 + v_2) + w_2, \ldots, (u_n + v_n) + w_n] \\ &= [u_1 + (v_1 + w_1), u_2 + (v_2 + w_2), \ldots, u_n + (v_n + w_n)] \\ &= [u_1, u_2, \ldots, u_n] + [v_1 + w_1, v_2 + w_2, \ldots, v_n + w_n] \\ &= [u_1, u_2, \ldots, u_n] + ([v_1, v_2, \ldots, v_n] + [w_1, w_2, \ldots, w_n]) \\ &= \mathbf{u} + (\mathbf{v} + \mathbf{w})\end{aligned}$$

Material com direitos autorais

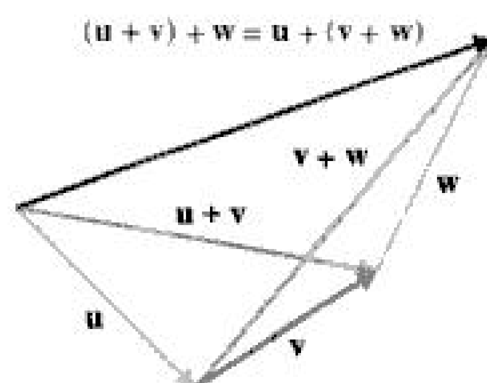


Figura 17

A quarta igualdade vale pela associatividade da adição de números reais. Note o uso cuidadoso dos parênteses. ◆

Pela propriedade (b) do Teorema 1, podemos, sem ambigüidade, escrever $\mathbf{u} + \mathbf{v} + \mathbf{w}$ omitindo os parênteses, já que podemos agrupar os somandos na ordem que quisermos. Por (a), também podemos rearranjar os somandos — por exemplo, como $\mathbf{w} + \mathbf{u} + \mathbf{v}$ — se assim o desejarmos. Da mesma maneira, somas de quatro ou mais vetores podem ser efetuadas sem levarmos em conta a ordem ou a maneira de agrupar. Em geral, se $\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_k$ são vetores de $\mathbb{R}^n$, escrevemos sua soma sem parênteses:

$$\mathbf{v}_1 + \mathbf{v}_2 + \cdots + \mathbf{v}_k$$

O próximo exemplo ilustra o uso do Teorema 1 ao efetuarmos cálculos algébricos com vetores.

EXEMPLO 5 Suponhamos que **a**, **b** e **x** sejam vetores em $\mathbb{R}^n$.

(a) Simplifique $3\mathbf{a} + (5\mathbf{b} - 2\mathbf{a}) + 2(\mathbf{b} - \mathbf{a})$.
(b) Se $5\mathbf{x} - \mathbf{a} = 2(\mathbf{a} + 2\mathbf{x})$, encontre **x** em função de **a**.

SOLUÇÃO: Daremos ambas as soluções com detalhes, fazendo referência a cada propriedade do Teorema 1 que usarmos. É uma boa prática justificar todos os passos nas primeiras vezes que você fizer esse tipo de cálculo. Uma vez familiarizado com as propriedades dos vetores, é aceitável que você omita alguns dos passos intermediários para economizar tempo e espaço.

(a) Começamos inserindo parênteses.

$$\begin{aligned}
3\mathbf{a} + (5\mathbf{b} - 2\mathbf{a}) + 2(\mathbf{b} - \mathbf{a}) &= (3\mathbf{a} + (5\mathbf{b} - 2\mathbf{a})) + 2(\mathbf{b} - \mathbf{a}) \\
&= (3\mathbf{a} + (-2\mathbf{a} + 5\mathbf{b})) + (2\mathbf{b} - 2\mathbf{a}) && \text{(a), (e)} \\
&= ((3\mathbf{a} + (-2\mathbf{a})) + 5\mathbf{b}) + (2\mathbf{b} - 2\mathbf{a}) && \text{(b)} \\
&= ((3 + (-2))\mathbf{a} + 5\mathbf{b}) + (2\mathbf{b} - 2\mathbf{a}) && \text{(f)} \\
&= (1\mathbf{a} + 5\mathbf{b}) + (2\mathbf{b} - 2\mathbf{a}) \\
&= ((\mathbf{a} + 5\mathbf{b}) + 2\mathbf{b}) - 2\mathbf{a} && \text{(b), (h)} \\
&= (\mathbf{a} + (5\mathbf{b} + 2\mathbf{b})) - 2\mathbf{a} && \text{(b)} \\
&= (\mathbf{a} + (5 + 2)\mathbf{b}) - 2\mathbf{a} && \text{(f)} \\
&= (7\mathbf{b} + \mathbf{a}) - 2\mathbf{a} && \text{(a)} \\
&= 7\mathbf{b} + (\mathbf{a} - 2\mathbf{a}) && \text{(b)} \\
&= 7\mathbf{b} + (1 - 2)\mathbf{a} && \text{(f), (h)} \\
&= 7\mathbf{b} + (-1)\mathbf{a} \\
&= 7\mathbf{b} - \mathbf{a}
\end{aligned}$$

Material com direitos autorais

Agora você pode ver por que concordaremos em omitir algumas dessas etapas! Na prática, é aceitável simplificar essa seqüência de operações como

$$\begin{aligned} 3\mathbf{a} + (5\mathbf{b} - 2\mathbf{a}) + 2(\mathbf{b} - \mathbf{a}) &= 3\mathbf{a} + 5\mathbf{b} - 2\mathbf{a} + 2\mathbf{b} - 2\mathbf{a} \\ &= (3\mathbf{a} - 2\mathbf{a} - 2\mathbf{a}) + (5\mathbf{b} + 2\mathbf{b}) \\ &= -\mathbf{a} + 7\mathbf{b} \end{aligned}$$

ou até fazer a maior parte dos cálculos mentalmente.

(b) Com detalhes, temos:

$$\begin{aligned} 5\mathbf{x} - \mathbf{a} &= 2(\mathbf{a} + 2\mathbf{x}) \\ 5\mathbf{x} - \mathbf{a} &= 2\mathbf{a} + 2(2\mathbf{x}) && \text{(e)} \\ 5\mathbf{x} - \mathbf{a} &= 2\mathbf{a} + (2 \cdot 2)\mathbf{x} && \text{(g)} \\ 5\mathbf{x} - \mathbf{a} &= 2\mathbf{a} + 4\mathbf{x} \\ (5\mathbf{x} - \mathbf{a}) - 4\mathbf{x} &= (2\mathbf{a} + 4\mathbf{x}) - 4\mathbf{x} \\ (-\mathbf{a} + 5\mathbf{x}) - 4\mathbf{x} &= 2\mathbf{a} + (4\mathbf{x} - 4\mathbf{x}) && \text{(a), (b)} \\ -\mathbf{a} + (5\mathbf{x} - 4\mathbf{x}) &= 2\mathbf{a} + \mathbf{0} && \text{(b), (d)} \\ -\mathbf{a} + (5 - 4)\mathbf{x} &= 2\mathbf{a} && \text{(f), (c)} \\ -\mathbf{a} + (1)\mathbf{x} &= 2\mathbf{a} \\ \mathbf{a} + (-\mathbf{a} + \mathbf{x}) &= \mathbf{a} + 2\mathbf{a} && \text{(h)} \\ (\mathbf{a} + (-\mathbf{a})) + \mathbf{x} &= (1 + 2)\mathbf{a} && \text{(b), (f)} \\ \mathbf{0} + \mathbf{x} &= 3\mathbf{a} && \text{(d)} \\ \mathbf{x} &= 3\mathbf{a} && \text{(c)} \end{aligned}$$

Novamente, omitiremos muitas dessas etapas na maioria dos casos. ◆

Combinações Lineares e Coordenadas

Um vetor que é uma soma de múltiplos escalares de outros vetores é chamado de *combinação linear* desses vetores. A definição formal é a seguinte:

Definição Um vetor $\mathbf{v}$ é uma ***combinação linear*** de vetores $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k$ se existirem escalares $c_1, c_2, \ldots, c_k$ tais que $\mathbf{v} = c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \ldots + c_k\mathbf{v}_k$. Os escalares $c_1, c_2, \ldots, c_k$ são chamados ***coeficientes*** da combinação linear.

EXEMPLO 6 O vetor $\begin{bmatrix} 2 \\ -2 \\ -1 \end{bmatrix}$ é uma combinação linear de $\begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}$, $\begin{bmatrix} 2 \\ -3 \\ 1 \end{bmatrix}$, e $\begin{bmatrix} 5 \\ -4 \\ 0 \end{bmatrix}$ pois

$$3\begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix} + 2\begin{bmatrix} 2 \\ -3 \\ 1 \end{bmatrix} - \begin{bmatrix} 5 \\ -4 \\ 0 \end{bmatrix} = \begin{bmatrix} 2 \\ -2 \\ -1 \end{bmatrix}$$ ◆

Material com direitos autorais

Observação: determinar se um dado vetor é combinação linear de outros vetores é um problema do qual nos ocuparemos no Capítulo 2.

Em $\mathbb{R}^2$, é possível desenhar combinações lineares de dois vetores (não paralelos) facilmente.

EXEMPLO 7 Sejam $\mathbf{u} = \begin{bmatrix} 3 \\ 1 \end{bmatrix}$ e $\mathbf{v} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$. Podemos utilizar **u** e **v** para constituir um novo sistema de eixos (do mesmo modo que $\mathbf{e}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ e $\mathbf{e}_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$ determinam o sistema de eixos usual). Com esses novos eixos, podemos obter uma ***rede coordenada*** que nos permita facilmente desenhar combinações lineares de **u** e **v**.

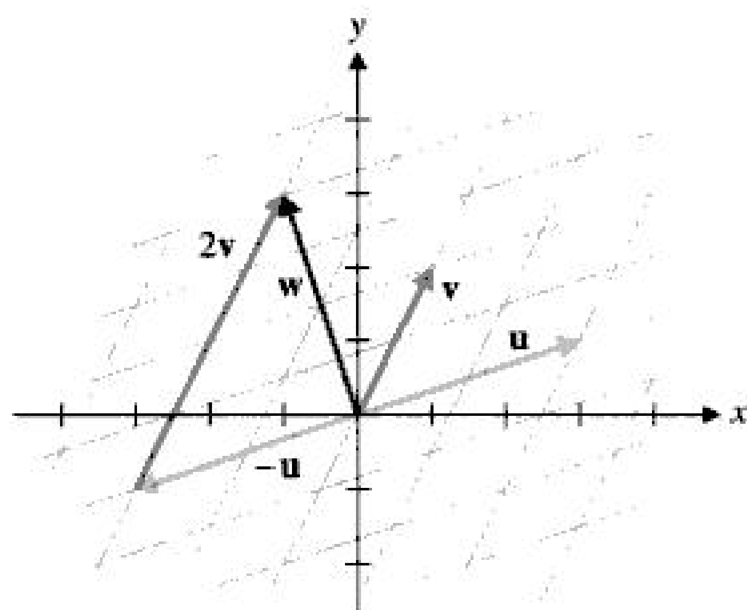


Figura 18

Como mostra a Figura 18, **w** pode ser localizado partindo-se da origem e percorrendo-se −**u** seguido de 2**v**. Isto é,

$$\mathbf{w} = -\mathbf{u} + 2\mathbf{v}$$

Dizemos que as coordenadas de **w** em relação a **u** e a **v** são −1 e 2. (Note que essa é apenas outra maneira de interpretar os coeficientes de uma combinação linear.) Segue que

$$\mathbf{w} = -\begin{bmatrix} 3 \\ 1 \end{bmatrix} + 2\begin{bmatrix} 1 \\ 2 \end{bmatrix} = \begin{bmatrix} -1 \\ 3 \end{bmatrix}$$

(Observe que −1 e 3 são as coordenadas de **w** em relação a $\mathbf{e}_1$ e $\mathbf{e}_2$.) ◆

Mudar do sistema de eixos usual para um outro alternativo é uma idéia útil. Tem aplicações em química e geologia, já que as estruturas moleculares e cristalinas freqüentemente não se adaptam à rede retangular. Neste livro, encontraremos essa idéia muitas vezes.

Material com direitos autorais

◆ EXERCÍCIOS 1.2 ◆

1. Desenhe os seguintes vetores em posição padrão em $\mathbb{R}^2$:

(a) $\mathbf{a} = \begin{bmatrix} 3 \\ 0 \end{bmatrix}$ (b) $\mathbf{b} = \begin{bmatrix} 2 \\ 3 \end{bmatrix}$

(c) $\mathbf{c} = \begin{bmatrix} -2 \\ 3 \end{bmatrix}$ (d) $\mathbf{d} = \begin{bmatrix} 3 \\ -2 \end{bmatrix}$

2. Desenhe os vetores do Exercício 1 com suas origens no ponto $(-2, -3)$.

3. Desenhe os seguintes vetores na posição padrão em $\mathbb{R}^3$:

(a) $\mathbf{a} = [0, 2, 0]$ (b) $\mathbf{b} = [3, 2, 1]$

(c) $\mathbf{c} = [1, -2, 1]$ (d) $\mathbf{d} = [-1, -1, -2]$

4. Se os vetores do Exercício 3 forem transladados de modo que suas extremidades estejam no ponto (4, 5, 6), ache os pontos correspondentes às suas origens.

5. Para cada um dos seguintes pares de pontos, desenhe o vetor $\overrightarrow{AB}$. Depois, determine e redesenhe $\overrightarrow{AB}$ na posição padrão.

(a) $A = (1, -1), B = (4, 2)$ (b) $A = (0, -2), B = (2, -1)$

(c) $A = (2, \frac{3}{2}), B = (\frac{1}{2}, 3)$ (d) $A = (\frac{1}{3}, \frac{1}{3}), B = (\frac{1}{6}, \frac{1}{2})$

6. Um excursionista anda 4 km na direção norte e depois 5 km na direção nordeste. Desenhe os vetores deslocamento que representam o passeio do excursionista e o vetor que representa o deslocamento a partir do ponto inicial.

Os Exercícios 7 a 10 se referem aos vetores do Exercício 1. Determine os vetores indicados e mostre como os resultados podem ser obtidos geometricamente.

7. $\mathbf{a} + \mathbf{b}$ **8.** $\mathbf{b} + \mathbf{c}$

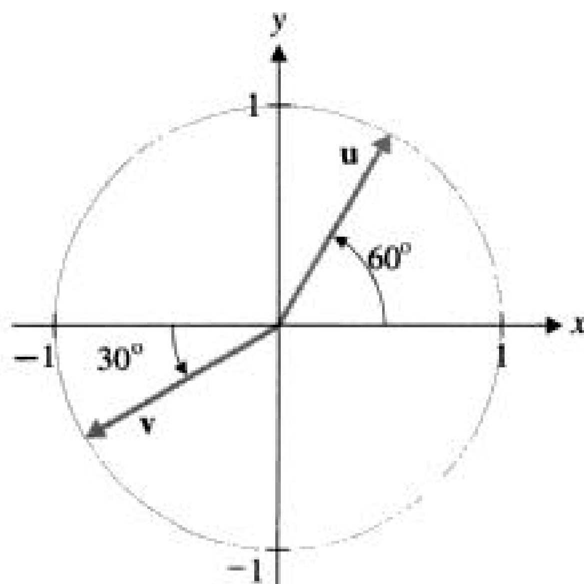
9. $\mathbf{d} - \mathbf{c}$ **10.** $\mathbf{a} - \mathbf{d}$

Os Exercícios 11 e 12 se referem aos vetores do Exercício 3. Determine os vetores indicados.

11. $2\mathbf{a} + 3\mathbf{c}$ **12.** $2\mathbf{c} - 3\mathbf{b} - \mathbf{d}$

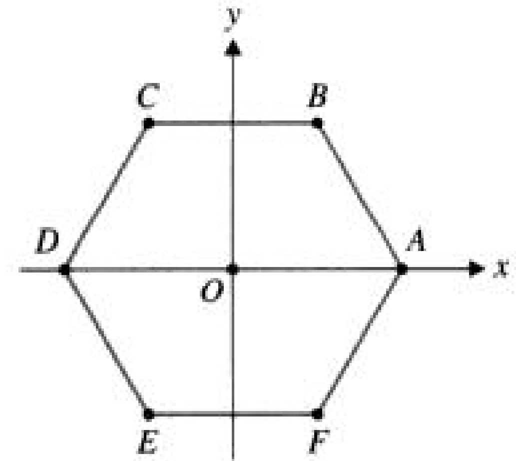
13. Ache as componentes dos vetores $\mathbf{u}$, $\mathbf{v}$, $\mathbf{u} + \mathbf{v}$ e $\mathbf{u} - \mathbf{v}$, onde $\mathbf{u}$ e $\mathbf{v}$ aparecem na Figura 19.

Figura 19



14. Na Figura 20, A, B, C, D, E e F são vértices de um hexágono regular centrado na origem.

Figura 20



Expresse cada um dos seguintes vetores em termos de $\mathbf{a} = \overrightarrow{OA}$ e $\mathbf{b} = \overrightarrow{OB}$:

(a) $\overrightarrow{AB}$ (b) $\overrightarrow{BC}$ (c) $\overrightarrow{AD}$

(d) $\overrightarrow{CF}$ (e) $\overrightarrow{AC}$

(f) $\overrightarrow{BC} + \overrightarrow{DE} + \overrightarrow{FA}$

Nos Exercícios 15 e 16, simplifique a expressão vetorial dada. Indique quais propriedades do Teorema 1 você usou.

15. $2(\mathbf{a} - 3\mathbf{b}) + 3(2\mathbf{b} + \mathbf{a})$

16. $-3(\mathbf{a} - \mathbf{c}) + 2(\mathbf{a} + 2\mathbf{b}) + 3(\mathbf{c} - \mathbf{b})$

Nos Exercícios 17 e 18, encontre o vetor $\mathbf{x}$ *em termos dos vetores* $\mathbf{a}$ *e* $\mathbf{b}$.

17. $\mathbf{x} - \mathbf{a} = 2(\mathbf{x} - 2\mathbf{a})$

18. $\mathbf{x} + 2\mathbf{a} - \mathbf{b} = 3(\mathbf{x} + \mathbf{a}) - 2(2\mathbf{a} - \mathbf{b})$

Nos Exercícios 19 e 20, desenhe os eixos coordenados relativos a $\mathbf{u}$ *e* $\mathbf{v}$ *e localize* $\mathbf{w}$.

19. $\mathbf{u} = \begin{bmatrix} 1 \\ -1 \end{bmatrix}, \mathbf{v} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \mathbf{w} = 2\mathbf{u} + 3\mathbf{v}$

20. $\mathbf{u} = \begin{bmatrix} -2 \\ 3 \end{bmatrix}, \mathbf{v} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}, \mathbf{w} = -\mathbf{u} - 2\mathbf{v}$

Nos Exercícios 21 e 22, desenhe os eixos coordenados usuais no mesmo diagrama que os eixos relativos a $\mathbf{u}$ *e* $\mathbf{v}$. *Use estes últimos para obter* $\mathbf{w}$ *como combinação linear de* $\mathbf{u}$ *e* $\mathbf{v}$.

21. $\mathbf{u} = \begin{bmatrix} 1 \\ -1 \end{bmatrix}, \mathbf{v} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \mathbf{w} = \begin{bmatrix} 2 \\ 6 \end{bmatrix}$

22. $\mathbf{u} = \begin{bmatrix} -2 \\ 3 \end{bmatrix}, \mathbf{v} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}, \mathbf{w} = \begin{bmatrix} 2 \\ 9 \end{bmatrix}$

23. Desenhe diagramas para ilustrar as propriedades (d) e (e) do Teorema 1.

24. Escreva demonstrações algébricas das propriedades (d) a (g) do Teorema 1.

Material com direitos autorais

1.3 Comprimento e Ângulo: O Produto Escalar

É muito fácil reformular os conceitos geométricos familiares de comprimento, distância e ângulo em termos de vetores. Tal procedimento permitirá que utilizemos essas idéias importantes e poderosas em ambientes mais gerais do que $\mathbb{R}^2$ e $\mathbb{R}^3$. Nos capítulos subseqüentes, elas serão usadas como ferramentas geométricas simples para resolver uma grande variedade de problemas que aparecerão nas aplicações – mesmo quando não houver nenhuma geometria aparente!

O Produto Escalar

As versões vetoriais de comprimento, distância e ângulo podem ser todas descritas usando-se a noção de produto escalar de dois vetores.

Definição Se

$$\mathbf{u} = \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix} \quad \text{e} \quad \mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix}$$

então o ***produto escalar*** $\mathbf{u} \cdot \mathbf{v}$ de $\mathbf{u}$ e $\mathbf{v}$ é definido por

$$\mathbf{u} \cdot \mathbf{v} = u_1v_1 + u_2v_2 + \cdots + u_nv_n$$

Em palavras, $\mathbf{u} \cdot \mathbf{v}$ é a soma dos produtos das componentes correspondentes de $\mathbf{u}$ e $\mathbf{v}$. É importante notar dois aspectos sobre esse "produto" que acabamos de definir. Primeiro, $\mathbf{u}$ e $\mathbf{v}$ devem ter o mesmo número de componentes; segundo, o produto escalar $\mathbf{u} \cdot \mathbf{v}$ é um *número*, não um vetor. (Este é o motivo pelo qual $\mathbf{u} \cdot \mathbf{v}$ chama-se ***produto escalar*** de $\mathbf{u}$ e $\mathbf{v}$.) O produto escalar de vetores em $\mathbb{R}^n$ é um caso especial e importante da noção mais geral de ***produto interno***, que será explorada no Capítulo 7.

EXEMPLO 1 Calcule $\mathbf{u} \cdot \mathbf{v}$, onde $\mathbf{u} = \begin{bmatrix} 1 \\ 2 \\ -3 \end{bmatrix}$ e $\mathbf{v} = \begin{bmatrix} -3 \\ 5 \\ 2 \end{bmatrix}$.

SOLUÇÃO: $\mathbf{u} \cdot \mathbf{v} = 1 \cdot (-3) + 2 \cdot 5 + (-3) \cdot 2 = 1$ ◆

Observe que, se tivéssemos calculado $\mathbf{v} \cdot \mathbf{u}$ no Exemplo 1, teríamos obtido

$$\mathbf{v} \cdot \mathbf{u} = (-3) \cdot 1 + 5 \cdot 2 + 2 \cdot (-3) = 1$$

Fica claro que $\mathbf{u} \cdot \mathbf{v} = \mathbf{v} \cdot \mathbf{u}$, já que os produtos individuais das coordenadas comutam. A comutatividade é uma das propriedades do produto escalar que usaremos com freqüência. As principais propriedades do produto escalar estão resumidas no Teorema 1.

◆ **TEOREMA 1** Sejam $\mathbf{u}$, $\mathbf{v}$ e $\mathbf{w}$ vetores de $\mathbb{R}^n$ e a um escalar. Então:

a. $\mathbf{u} \cdot \mathbf{v} = \mathbf{v} \cdot \mathbf{u}$ — Comutatividade
b. $\mathbf{u} \cdot (\mathbf{v} + \mathbf{w}) = \mathbf{u} \cdot \mathbf{v} + \mathbf{u} \cdot \mathbf{w}$ — Distributividade
c. $a\mathbf{u} \cdot \mathbf{v} = a(\mathbf{u} \cdot \mathbf{v})$
d. $\mathbf{u} \cdot \mathbf{u} \geq 0$ e $\mathbf{u} \cdot \mathbf{u} = 0$ se, e somente se, $\mathbf{u} = \mathbf{0}$.

Material com direitos autorais

DEMONSTRAÇÃO: Provaremos (a) e (c) e deixaremos a demonstração das demais propriedades para os exercícios. (a) Aplicando a definição de produto escalar para $\mathbf{u} \cdot \mathbf{v}$ e $\mathbf{v} \cdot \mathbf{u}$, obtemos

$$\begin{aligned}\mathbf{u} \cdot \mathbf{v} &= u_1v_1 + u_2v_2 + \dots + u_nv_n \\ &= v_1u_1 + v_2u_2 + \dots + v_nu_n \\ &= \mathbf{v} \cdot \mathbf{u}\end{aligned}$$

onde a igualdade do meio segue do fato de a multiplicação de números reais ser comutativa.

(c) Usando as definições de produto de números e produto escalar, temos

$$\begin{aligned}(a\mathbf{u}) \cdot \mathbf{v} &= [au_1, au_2, \dots, au_n] \cdot [v_1, v_2, \dots, v_n] \\ &= au_1v_1 + au_2v_2 + \dots + au_nv_n \\ &= a(u_1v_1 + u_2v_2 + \dots + u_nv_n) \\ &= a(\mathbf{u} \cdot \mathbf{v})\end{aligned}$$

◆

Observações:

- A propriedade (b) pode ser lida da direita para a esquerda, e, neste caso, diz que podemos pôr o vetor $\mathbf{u}$ em evidência na soma dos produtos escalares. Também vale a propriedade análoga, do "lado direito" — que decorre das propriedades (b) e (a): $(\mathbf{v} + \mathbf{w}) \cdot \mathbf{u} = \mathbf{v} \cdot \mathbf{u} + \mathbf{w} \cdot \mathbf{u}$.
- A propriedade (c) pode ser estendida para obtermos $\mathbf{u} \cdot (a\mathbf{v}) = a(\mathbf{u} \cdot \mathbf{v})$ (Exercício 44). Tal versão estendida de (c) diz essencialmente que, tomando-se um múltiplo escalar de um produto escalar de vetores, o escalar pode ser combinado com o vetor mais conveniente. Por exemplo:

$$(\tfrac{1}{2}[-1, -3, 2]) \cdot [6, -4, 0] = [-1, -3, 2] \cdot (\tfrac{1}{2}[6, -4, 0]) = [-1, -3, 2] \cdot [3, -2, 0] = 3$$

Com essa escolha, evitamos introduzir frações entre os vetores, como teria o agrupamento original.

- A segunda parte de (d) usa o conectivo lógico *se, e somente se*. O Apêndice A discute essa expressão com mais detalhe, mas, para o momento, vamos apenas observar que ela indica uma *dupla implicação* – isto é,

$$\text{se } \mathbf{u} = \mathbf{0}, \text{ então } \mathbf{u} \cdot \mathbf{u} = 0$$

e

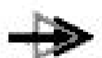
$$\text{se } \mathbf{u} \cdot \mathbf{u} = 0, \text{ então } \mathbf{u} = \mathbf{0}$$

O Teorema 1 mostra que aspectos da álgebra dos vetores lembram a álgebra dos números. O próximo exemplo mostra que algumas vezes podemos encontrar análogos vetoriais de identidades familiares.

EXEMPLO 2 Prove que $(\mathbf{u} + \mathbf{v}) \cdot (\mathbf{u} + \mathbf{v}) = \mathbf{u} \cdot \mathbf{u} + 2(\mathbf{u} \cdot \mathbf{v}) + \mathbf{v} \cdot \mathbf{v}$ para quaisquer vetores $\mathbf{u}$ e $\mathbf{v}$ de $\mathbb{R}^n$.

SOLUÇÃO:

$$\begin{aligned}(\mathbf{u} + \mathbf{v}) \cdot (\mathbf{u} + \mathbf{v}) &= (\mathbf{u} + \mathbf{v}) \cdot \mathbf{u} + (\mathbf{u} + \mathbf{v}) \cdot \mathbf{v} \\ &= \mathbf{u} \cdot \mathbf{u} + \mathbf{v} \cdot \mathbf{u} + \mathbf{u} \cdot \mathbf{v} + \mathbf{v} \cdot \mathbf{v} \\ &= \mathbf{u} \cdot \mathbf{u} + \mathbf{u} \cdot \mathbf{v} + \mathbf{u} \cdot \mathbf{v} + \mathbf{v} \cdot \mathbf{v} \\ &= \mathbf{u} \cdot \mathbf{u} + 2(\mathbf{u} \cdot \mathbf{v}) + \mathbf{v} \cdot \mathbf{v}\end{aligned}$$

 (Identifique as propriedades do Teorema 1 que foram usadas em cada etapa.) ◆

Material com direitos autorais

Comprimento

Para entender qual a relação entre o produto escalar e o cálculo de comprimentos, recordemos como comprimentos são calculados no plano. O Teorema de Pitágoras é tudo de que precisamos.

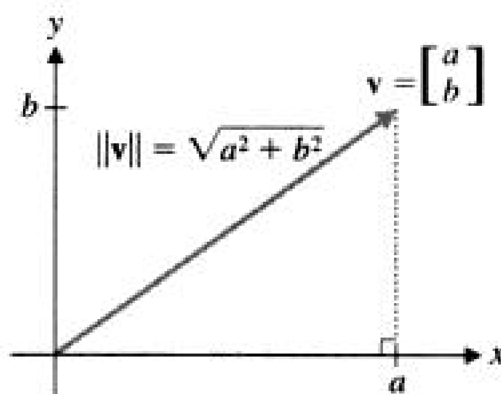


Figura 1

Em $\mathbb{R}^2$, o comprimento do vetor $\mathbf{v} = \begin{bmatrix} a \\ b \end{bmatrix}$ é a distância da origem ao ponto (a, b), a qual, pelo Teorema de Pitágoras, é dada por $\sqrt{a^2 + b^2}$, como na Figura 1. Observe que $a^2 + b^2 = \mathbf{v} \cdot \mathbf{v}$. Isso leva à seguinte definição:

Definição O ***comprimento*** (ou ***norma***) de um vetor $\mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix}$ em $\mathbb{R}^n$ é o *número* não **negativo** $\|\mathbf{v}\|$ definido por

$$\|\mathbf{v}\| = \sqrt{\mathbf{v} \cdot \mathbf{v}} = \sqrt{v_1^2 + v_2^2 + \cdots + v_n^2}$$

Em palavras, o comprimento de um vetor é a raiz quadrada da soma dos quadrados de suas coordenadas. Observe que a raiz quadrada de $\mathbf{v} \cdot \mathbf{v}$ está sempre definida, pois $\mathbf{v} \cdot \mathbf{v} \geq 0$, pelo Teorema 1(d). Note também que a definição pode ser reescrita como $\|\mathbf{v}\|^2 = \mathbf{v} \cdot \mathbf{v}$. Isso será de grande utilidade na demonstração de outras propriedades do produto escalar e do comprimento de vetores.

EXEMPLO 3 $\|[2, 3]\| = \sqrt{2^2 + 3^2} = \sqrt{13}$ ◆

O Teorema 2 enumera algumas das principais propriedades do comprimento de um vetor:

◆ **TEOREMA 2** Seja $\mathbf{v}$ um vetor em $\mathbb{R}^n$ e c um escalar. Então:

a. $\|\mathbf{v}\| = 0$ se, e somente se, $\mathbf{v} = \mathbf{0}$
b. $\|c\mathbf{v}\| = |c|\,\|\mathbf{v}\|$

DEMONSTRAÇÃO: A propriedade (a) segue imediatamente do Teorema 1(d). Para mostrar (b), temos

$$\|c\mathbf{v}\|^2 = (c\mathbf{v}) \cdot (c\mathbf{v}) = c^2(\mathbf{v} \cdot \mathbf{v}) = c^2\|\mathbf{v}\|^2$$

usando o Teorema 1(c). Tomando-se a raiz quadrada de ambos os lados, e usando o fato de que $\sqrt{c^2} = |c|$, para todo número real c, obtemos o resultado. ◆

Um vetor de comprimento 1 é chamado ***vetor unitário***. Em $\mathbb{R}^2$, o conjunto de todos os vetores unitários pode ser identificado com o *círculo unitário*, o círculo de raio 1 centrado na origem (veja a Figura 2). Dado qualquer vetor não nulo $\mathbf{v}$, podemos sempre encontrar um vetor unitário, de mesma direção e sentido que $\mathbf{v}$,

Material com direitos autorais

dividindo **v** por seu próprio comprimento (ou, equivalentemente, *multiplicando* por $1/\|\mathbf{v}\|$). Podemos mostrar esse fato algebricamente usando a propriedade (b) do Teorema 2: se $\mathbf{u} = (1/\|\mathbf{v}\|)\mathbf{v}$, então

$$\|\mathbf{u}\| = \|(1/\|\mathbf{v}\|)\mathbf{v}\| = |1/\|\mathbf{v}\|| \, \|\mathbf{v}\| = (1/\|\mathbf{v}\|)\|\mathbf{v}\| = 1$$

e **u** possui a mesma direção e sentido que **v**, pois $1/\|\mathbf{v}\|$ é um número positivo. Achar um vetor unitário de mesma direção e sentido é, em geral, um processo ao qual nos referimos como ***normalizar*** um vetor[5] (veja a Figura 3).

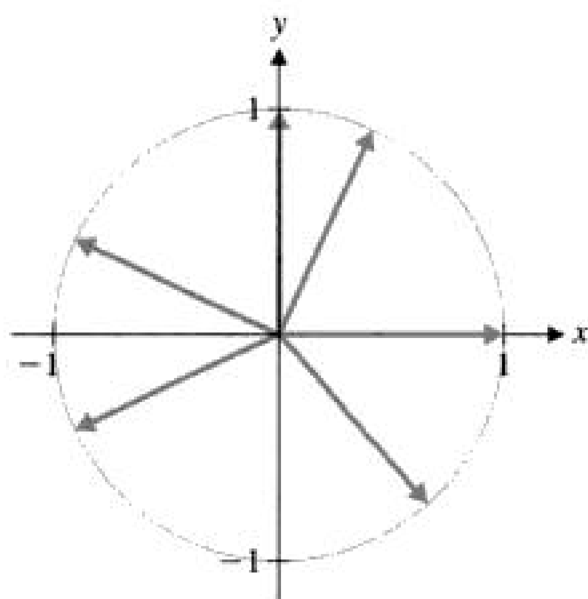


Figura 2 Vetores unitários em $\mathbb{R}^2$

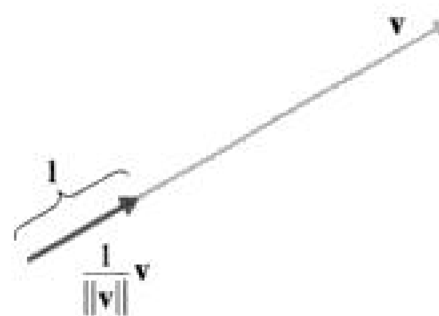


Figura 3 Normalização de um vetor

EXEMPLO 4 Em $\mathbb{R}^2$, sejam $\mathbf{e}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ e $\mathbf{e}_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$. Então, $\mathbf{e}_1$ e $\mathbf{e}_2$ são vetores unitários, pois a soma dos quadrados de suas componentes, em cada caso, é 1. Analogamente, em $\mathbb{R}^3$, podemos construir vetores unitários

$$\mathbf{e}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \quad \mathbf{e}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \quad \text{e} \quad \mathbf{e}_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

Observe na Figura 4, que esses vetores ajudam a localizar os eixos de coordenadas positivas em $\mathbb{R}^2$ e $\mathbb{R}^3$. ◆

Em geral, em $\mathbb{R}^n$, definimos vetores unitários $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n$, onde $\mathbf{e}_i$ possui 1 na i-ésima coordenada, e zero, nas demais. Tais vetores aparecem com freqüência na álgebra linear e são chamados de ***vetores unitários canônicos***.

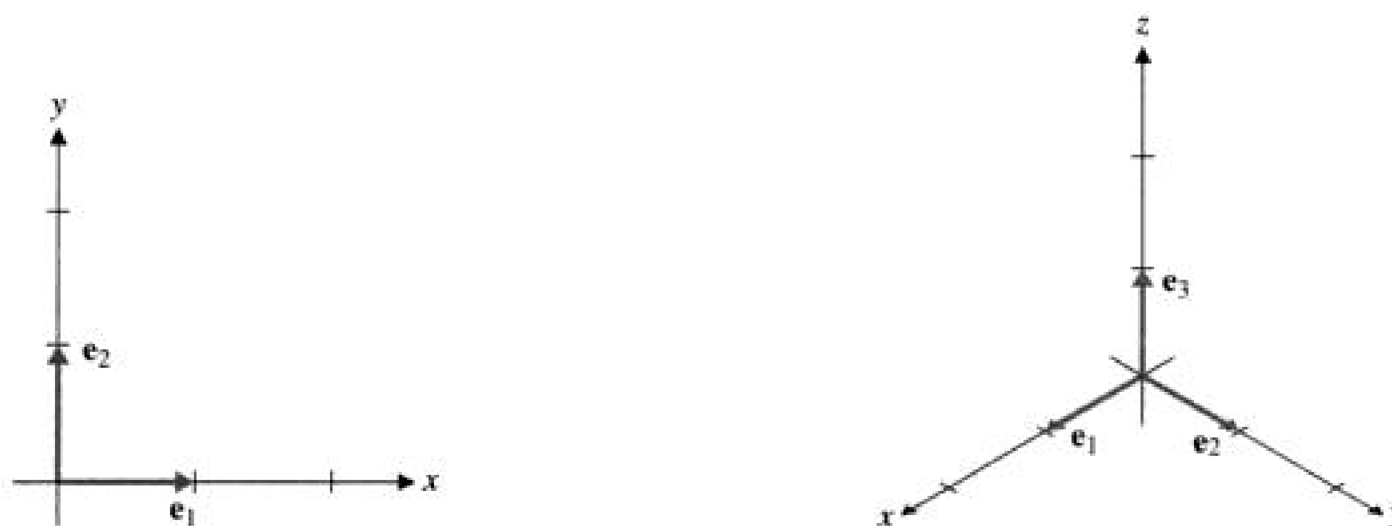


Figura 4 Vetores unitários canônicos em $\mathbb{R}^2$ e $\mathbb{R}^3$

[5] N.T.: Em português, chamamos *versor de* **v** ao vetor unitário $\mathbf{u} = \frac{1}{\|\mathbf{v}\|}\mathbf{v}$.

Material com direitos autorais

EXEMPLO 5 Normalize o vetor $\mathbf{v} = \begin{bmatrix} 2 \\ -1 \\ 3 \end{bmatrix}$.

SOLUÇÃO: $\|\mathbf{v}\| = \sqrt{2^2 + (-1)^2 + 3^2} = \sqrt{14}$, portanto, um vetor unitário com mesma direção e sentido de $\mathbf{v}$ é dado por

$$\mathbf{u} = (1/\|\mathbf{v}\|)\mathbf{v} = (1/\sqrt{14})\begin{bmatrix} 2 \\ -1 \\ 3 \end{bmatrix} = \begin{bmatrix} 2/\sqrt{14} \\ -1/\sqrt{14} \\ 3/\sqrt{14} \end{bmatrix}$$ ◆

Como a propriedade (b) do Teorema 2 descreve de que forma o comprimento se comporta em relação à multiplicação por escalar, a curiosidade natural sugere que perguntemos se comprimento e adição de vetores são compatíveis. Seria muito bom se tivéssemos uma identidade tal que $\|\mathbf{u} + \mathbf{v}\| = \|\mathbf{u}\| + \|\mathbf{v}\|$, mas, para quase toda escolha de vetores $\mathbf{u}$ e $\mathbf{v}$, isso é falso. [Veja o Exercício 42(a).] No entanto, nem tudo está perdido: se trocarmos o sinal = por ≤, a desigualdade resultante será verdadeira. A demonstração desse famoso e importante resultado — a Desigualdade Triangular — depende de outra importante desigualdade: a Desigualdade de Cauchy-Schwarz, que demonstraremos e discutiremos mais detalhadamente no Capítulo 7.

◆ **TEOREMA 3 A Desigualdade de Cauchy-Schwarz**

Para quaisquer vetores $\mathbf{u}$ e $\mathbf{v}$ em $\mathbb{R}^n$,

$$|\mathbf{u} \cdot \mathbf{v}| \leq \|\mathbf{u}\|\|\mathbf{v}\|$$

Veja os Exercícios 55 e 56 para abordagens algébricas e geométricas das demonstrações dessa desigualdade.

Em $\mathbb{R}^2$ ou $\mathbb{R}^3$, onde podemos usar geometria, fica claro, a partir de um diagrama como o da Figura 5, que vale $\|\mathbf{u} + \mathbf{v}\| \leq \|\mathbf{u}\| + \|\mathbf{v}\|$, para quaisquer vetores $\mathbf{u}$ e $\mathbf{v}$. Mostraremos agora que isso é verdade em geral.

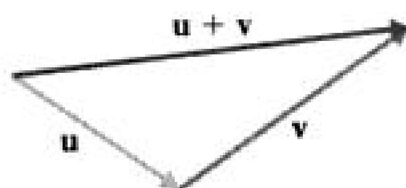


Figura 5 A Desigualdade Triangular

◆ **TEOREMA 4 A Desigualdade Triangular**

Para quaisquer vetores $\mathbf{u}$ e $\mathbf{v}$ em $\mathbb{R}^n$,

$$\|\mathbf{u} + \mathbf{v}\| \leq \|\mathbf{u}\| + \|\mathbf{v}\|$$

DEMONSTRAÇÃO: Como ambos os lados da desigualdade são não negativos, mostrar que o *quadrado* do lado esquerdo é menor ou igual ao *quadrado* do lado direito é equivalente a provar



o teorema. (Por quê?) Calculando, obtemos:

$$\begin{aligned} \|\mathbf{u} + \mathbf{v}\|^2 &= (\mathbf{u} + \mathbf{v}) \cdot (\mathbf{u} + \mathbf{v}) \\ &= \mathbf{u} \cdot \mathbf{u} + 2(\mathbf{u} \cdot \mathbf{v}) + \mathbf{v} \cdot \mathbf{v} && \text{Pelo Exemplo 2} \\ &\leq \|\mathbf{u}\|^2 + 2|\mathbf{u} \cdot \mathbf{v}| + \|\mathbf{v}\|^2 \\ &\leq \|\mathbf{u}\|^2 + 2\|\mathbf{u}\|\|\mathbf{v}\| + \|\mathbf{v}\|^2 && \text{Por Cauchy-Schwarz} \\ &= (\|\mathbf{u}\| + \|\mathbf{v}\|)^2 \end{aligned}$$

como desejado. ◆

Material com direitos autorais

Distância

O conceito de distância entre dois vetores é análogo ao de distância entre dois pontos na reta real ou dois pontos no plano cartesiano. Na reta real (Figura 6), a distância entre os números a e b é dada por $|a - b|$. (Tomar o valor absoluto assegura que não precisemos nos preocupar qual, a ou b, é maior.) Essa distância é também igual a $\sqrt{(a-b)^2}$, e sua generalização para duas dimensões é a fórmula familiar para a distância d entre pontos (a_1, a_2) e (b_1, b_2) — a saber, $d = \sqrt{(a_1 - b_1)^2 + (a_2 - b_2)^2}$.

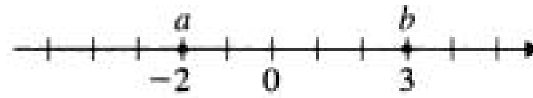
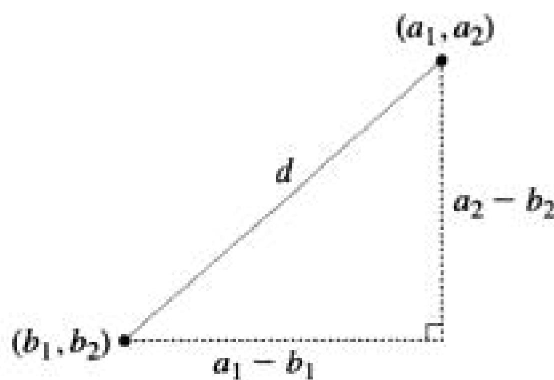


Figura 6 $d = |a - b| = |-2-3| = 5$

Em termos de vetores, se $\mathbf{a} = \begin{bmatrix} a_1 \\ a_2 \end{bmatrix}$ e $\mathbf{b} = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}$, então d é exatamente o comprimento de $\mathbf{a} - \mathbf{b}$, como mostra a Figura 7. Essa é a base para a próxima definição.



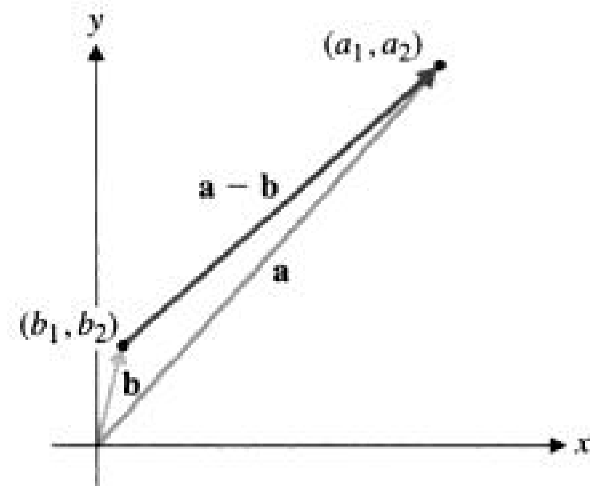


Figura 7 $d = \sqrt{(a_1 - b_1)^2 + (a_2 - b_2)^2} = \|\mathbf{a} - \mathbf{b}\|$

Definição A ***distância*** $d(\mathbf{u}, \mathbf{v})$ entre vetores $\mathbf{u}$ e $\mathbf{v}$, em $\mathbb{R}^n$, é definida por

$$d(\mathbf{u}, \mathbf{v}) = \|\mathbf{u} - \mathbf{v}\|$$

EXEMPLO 6 Determine a distância entre $\mathbf{u} = \begin{bmatrix} \sqrt{2} \\ 1 \\ -1 \end{bmatrix}$ e $\mathbf{v} = \begin{bmatrix} 0 \\ 2 \\ -2 \end{bmatrix}$.

SOLUÇÃO: Calculamos $\mathbf{u} - \mathbf{v} = \begin{bmatrix} \sqrt{2} \\ -1 \\ 1 \end{bmatrix}$, então

$$d(\mathbf{u}, \mathbf{v}) = \|\mathbf{u} - \mathbf{v}\| = \sqrt{(\sqrt{2})^2 + (-1)^2 + 1^2} = \sqrt{4} = 2$$



Material com direitos autorais

Ângulos

O produto escalar também pode ser usado para determinar o ângulo entre um par de vetores. Em $\mathbb{R}^2$ ou $\mathbb{R}^3$, o ângulo entre os vetores não nulos **u** e **v** será o ângulo θ, determinado por esses vetores, que satisfizer $0 \leq \theta \leq 180°$ (veja a Figura 8).

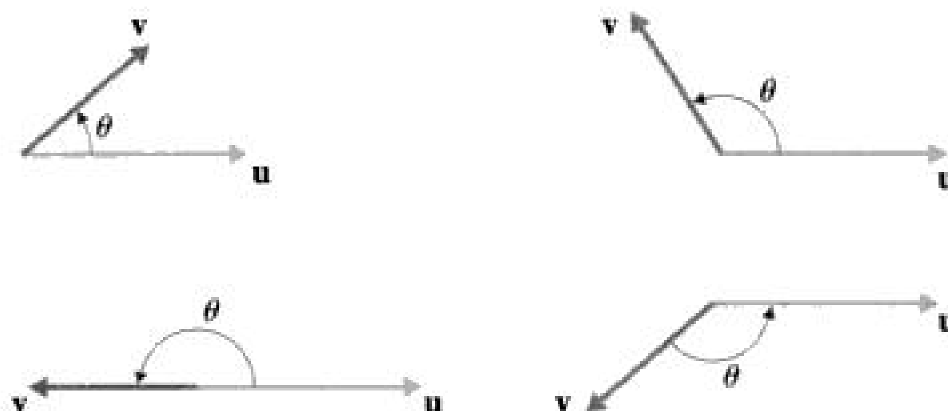


Figura 8 O ângulo entre **u** e **v**

Na Figura 9, considere o triângulo de lados **u**, **v** e **u** – **v**, onde θ é o ângulo entre **u** e **v**. Aplicando a lei dos co-senos a esse triângulo, vem que

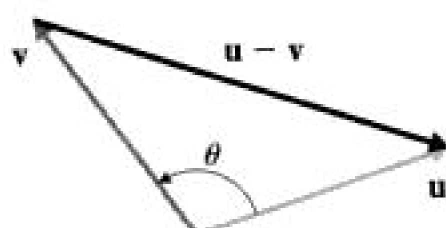


Figura 9

$$\|\mathbf{u} - \mathbf{v}\|^2 = \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2 - 2\|\mathbf{u}\|\,\|\mathbf{v}\| \cos\theta$$

Desenvolvendo o lado esquerdo e usando a propriedade $\|\mathbf{v}\|^2 = \mathbf{v} \cdot \mathbf{v}$ várias vezes, obtemos

$$\|\mathbf{u}\|^2 - 2(\mathbf{u} \cdot \mathbf{v}) + \|\mathbf{v}\|^2 = \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2 - 2\|\mathbf{u}\|\,\|\mathbf{v}\| \cos\theta$$

que, depois de algumas simplificações, nos leva a $\mathbf{u} \cdot \mathbf{v} = \|\mathbf{u}\|\,\|\mathbf{v}\| \cos\theta$. Dessa última expressão obtemos a seguinte fórmula para o ângulo θ entre vetores não nulos **u** e **v** em $\mathbb{R}^2$ ou $\mathbb{R}^3$:

$$\cos\theta = \frac{\mathbf{u} \cdot \mathbf{v}}{\|\mathbf{u}\|\,\|\mathbf{v}\|}$$

EXEMPLO 7 Determine o ângulo entre os vetores $\mathbf{u} = [2, 1, -2]$ e $\mathbf{v} = [1, 1, 1]$.

SOLUÇÃO: Calculamos $\mathbf{u} \cdot \mathbf{v} = 2 \cdot 1 + 1 \cdot 1 + (-2) \cdot 1 = 1$, $\|\mathbf{u}\| = \sqrt{2^2 + 1^2 + (-2)^2} = \sqrt{9} = 3$, e $\|\mathbf{v}\| = \sqrt{1^2 + 1^2 + 1^2} = \sqrt{3}$. Logo, $\cos\theta = 1/3\sqrt{3}$, e portanto, $\theta = \cos^{-1}(1/3\sqrt{3}) \approx 1{,}377$ radianos. ◆

EXEMPLO 8 Determine o ângulo entre as diagonais de duas faces adjacentes de um cubo.

Material com direitos autorais

SOLUÇÃO: As dimensões do cubo não importam. Assim, consideraremos um cubo com lados de comprimento 1. Escolha um sistema de eixos coordenados em $\mathbb{R}^3$ como mostra a Figura 10, e tome as diagonais representadas pelos vetores $[1, 0, 1]$ e $[0, 1, 1]$. O ângulo θ entre esses vetores satisfaz

$$\cos\theta = \frac{1\cdot 0 + 0\cdot 1 + 1\cdot 1}{\sqrt{2}\sqrt{2}} = \frac{1}{2}$$

de onde segue que o ângulo procurado é $\pi/3$ radianos, ou 60°.

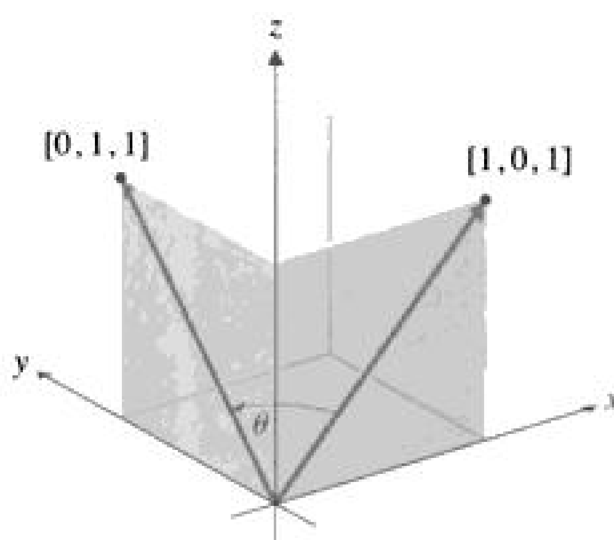


Figura 10

(Na verdade, não precisamos fazer nenhum cálculo para obter essa resposta. Se desenharmos uma terceira diagonal unindo os vértices (1, 0, 1) e (0, 1, 1), teremos um triângulo equilátero, já que todas as diagonais das faces têm o mesmo comprimento. O ângulo que procuramos é um dos ângulos desse triângulo, e, portanto, mede 60°. Às vezes, uma pequena idéia pode economizar muitos cálculos; em casos como este, ela fornece uma verificação confortável em nosso trabalho!) ◆

Observações:

- Como mostra esta discussão, geralmente teremos de estabelecer uma aproximação para o ângulo entre dois vetores. Entretanto, quando o ângulo for um dos chamados ângulos especiais (0°, 30°, 45°, 60°, 90°, ou um múltiplo inteiro desses), deveremos ser capazes de reconhecer seu co-seno (Tabela 1) e então obter o correspondente ângulo com exatidão. Em todos os outros casos, usaremos uma calculadora ou computador para aproximar o ângulo desejado por meio da função arco – co-seno.

Tabela 1 Co-senos de Ângulos Especiais

θ	0°	30°	45°	60°	90°
$\cos\theta$	$\frac{\sqrt{4}}{2} = 1$	$\frac{\sqrt{3}}{2}$	$\frac{\sqrt{2}}{2} = \frac{1}{\sqrt{2}}$	$\frac{\sqrt{1}}{2} = \frac{1}{2}$	$\frac{\sqrt{0}}{2} = 0$

- A dedução da fórmula para o co-seno do ângulo entre dois vetores é válida somente em $\mathbb{R}^2$ ou $\mathbb{R}^3$, já que depende de um fato geométrico: a lei dos co-senos. Em $\mathbb{R}^n$, para $n > 3$, a fórmula pode ser considerada uma *definição*. Isso faz sentido, pois a desigualdade de Cauchy-Schwarz implica que $\left|\frac{\mathbf{u}\cdot\mathbf{v}}{\|\mathbf{u}\|\,\|\mathbf{v}\|}\right| \leq 1$, e, portanto, $\frac{\mathbf{u}\cdot\mathbf{v}}{\|\mathbf{u}\|\,\|\mathbf{v}\|}$ varia de –1 a 1, exatamente como o co-seno.

Material com direitos autorais

Vetores Ortogonais

O conceito de perpendicularidade é fundamental em geometria. Qualquer pessoa que estuda geometria compreende rapidamente a importância e utilidade dos ângulos retos. Generalizaremos agora a idéia de perpendicularidade para vetores em $\mathbb{R}^n$, onde é chamada de ***ortogonalidade***.

Em $\mathbb{R}^2$ ou $\mathbb{R}^3$, dois vetores não nulos **u** e **v** são perpendiculares se o ângulo θ entre eles é um ângulo reto – isto é, se $\theta = \pi/2$ radianos, ou 90º.

Assim, $\frac{\mathbf{u}\cdot\mathbf{v}}{\|\mathbf{u}\|\,\|\mathbf{v}\|} = \cos 90^\circ = 0$, e, portanto, $\mathbf{u}\cdot\mathbf{v} = 0$.

Isso motiva a seguinte definição:

> A palavra ortogonal é derivada da palavra grega orthos, que significa "reto", e gonia, que significa "ângulo". Por isso, ortogonal significa literalmente "ângulo reto". O equivalente em latim é retangular.

Definição Dois vetores **u** e **v** em $\mathbb{R}^n$ são ***ortogonais*** entre si se $\mathbf{u}\cdot\mathbf{v} = 0$.

Como $\mathbf{0}\cdot\mathbf{v} = 0$, para todo vetor **v** em $\mathbb{R}^n$, dizemos que o vetor nulo é ortogonal a qualquer vetor.

EXEMPLO 9 Em $\mathbb{R}^3$, $\mathbf{u} = [1, 1, -2]$ e $\mathbf{v} = [3, 1, 2]$ são ortogonais, pois $\mathbf{u}\cdot\mathbf{v} = 3 + 1 - 4 = 0$. ◆

EXEMPLO 10 Do Exemplo 2, temos que $\|\mathbf{u} + \mathbf{v}\|^2 = \|\mathbf{u}\|^2 + 2(\mathbf{u}\cdot\mathbf{v}) + \|\mathbf{v}\|^2$, para quaisquer vetores **u** e **v** em $\mathbb{R}^n$. Segue imediatamente que $\|\mathbf{u} + \mathbf{v}\|^2 = \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2$ se, e somente se, $\mathbf{u}\cdot\mathbf{v} = 0$. Este é o Teorema de Pitágoras para $\mathbb{R}^n$ (Figura 11).

$\|\mathbf{u}+\mathbf{v}\|^2 = \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2$ se, e somente se, **u** e **v** são ortogonais. ◆

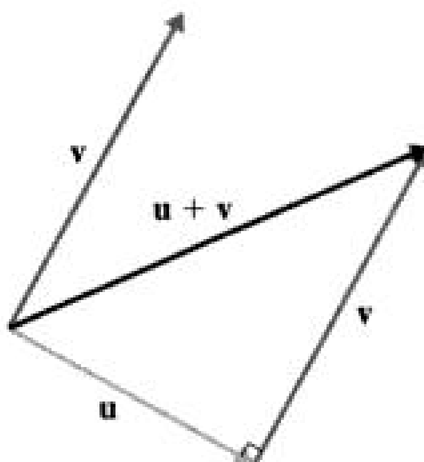


Figura 11 O Teorema de Pitágoras

O conceito de ortogonalidade é um dos mais importantes e úteis em álgebra linear, e freqüentemente aparece de maneiras surpreendentes. O Capítulo 5 oferece um tratamento detalhado deste tópico, mas iremos encontrá-lo muitas vezes antes disso. Um problema no qual a ortogonalidade claramente desempenha um papel importante é na determinação da distância de um ponto a uma reta, onde "baixar a perpendicular" é um procedimento familiar.

Material com direitos autorais

Projeções

Consideraremos agora o problema de encontrar a distância de um ponto a uma reta no contexto dos vetores. Como veremos, essa técnica conduz a um importante conceito: o da projeção de um vetor sobre outro.

Como mostra a Figura 12, o problema de achar a distância de um ponto B a uma reta r (em $\mathbb{R}^2$ ou $\mathbb{R}^3$) se reduz ao problema de achar o comprimento do segmento de reta $\overline{PB}$, perpendicular a r, ou, equivalentemente, o comprimento do vetor $\overrightarrow{PB}$. Se escolhermos um ponto A sobre r, então, no triângulo retângulo ΔAPB, os outros dois vetores são o cateto $\overrightarrow{AP}$ e a hipotenusa $\overrightarrow{AB}$. O vetor $\overrightarrow{AP}$ é chamado de *projeção* de $\overrightarrow{AB}$ sobre a reta r. Interpretaremos agora esta situação em termos de vetores.

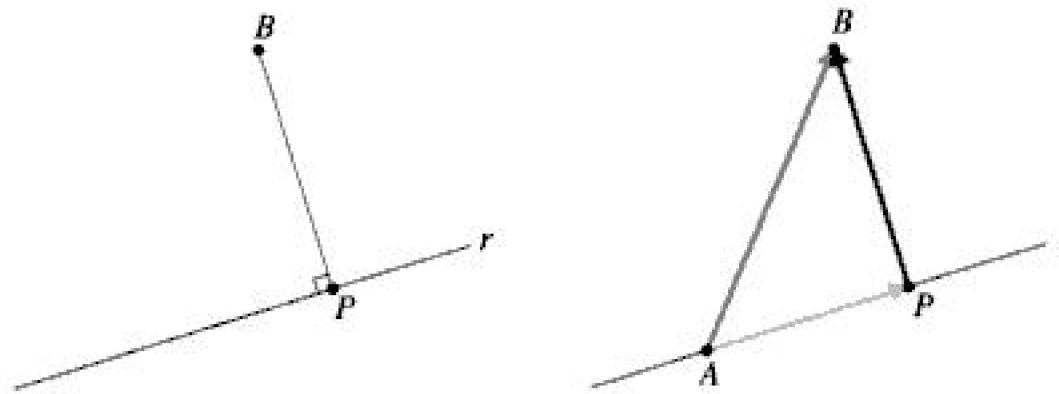


Figura 12 A distância de um ponto a uma reta

Considere dois vetores não nulos $\mathbf{u}$ e $\mathbf{v}$. Seja $\mathbf{p}$ o vetor obtido ao traçarmos a perpendicular a partir da extremidade de $\mathbf{v}$ sobre $\mathbf{u}$, e seja θ o ângulo entre $\mathbf{u}$ e $\mathbf{v}$, como mostra a Figura 13. Então, claramente $\mathbf{p} = \|\mathbf{p}\|\hat{\mathbf{u}}$, onde $\hat{\mathbf{u}} = (1/\|\mathbf{u}\|)\mathbf{u}$ é o versor de $\mathbf{u}$. Além disso, da trigonometria elementar temos $\|\mathbf{p}\| = \|\mathbf{v}\| \cos\theta$, e sabemos que $\cos\theta = \dfrac{\mathbf{u}\cdot\mathbf{v}}{\|\mathbf{u}\|\,\|\mathbf{v}\|}$. Dessa forma, depois de algumas substituições, obtemos:

$$\begin{aligned}\mathbf{p} &= \|\mathbf{v}\|\left(\frac{\mathbf{u}\cdot\mathbf{v}}{\|\mathbf{u}\|\,\|\mathbf{v}\|}\right)\left(\frac{1}{\|\mathbf{u}\|}\right)\mathbf{u}\\ &= \left(\frac{\mathbf{u}\cdot\mathbf{v}}{\|\mathbf{u}\|^2}\right)\mathbf{u}\\ &= \left(\frac{\mathbf{u}\cdot\mathbf{v}}{\mathbf{u}\cdot\mathbf{u}}\right)\mathbf{u}\end{aligned}$$

Essa é a fórmula que queremos, e é a motivação para a seguinte definição válida para vetores em $\mathbb{R}^n$:

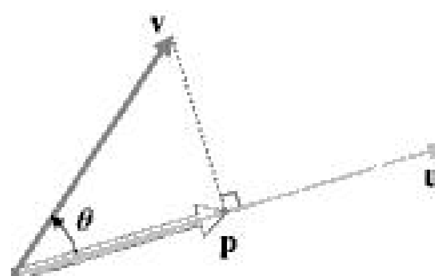


Figura 13 A projeção de $\mathbf{v}$ sobre $\mathbf{u}$

Material com direitos autorais

Definição Se **u** e **v** são vetores em $\mathbb{R}^n$ e $\mathbf{u} \neq \mathbf{0}$, a ***projeção de* v *sobre* u** é o vetor $\text{proj}_{\mathbf{u}}(\mathbf{v})$ definido por

$$\text{proj}_{\mathbf{u}}(\mathbf{v}) = \left(\frac{\mathbf{u} \cdot \mathbf{v}}{\mathbf{u} \cdot \mathbf{u}}\right)\mathbf{u}$$

Um modo alternativo de obter essa fórmula é descrito no Exercício 57.

Observações:

- O termo *projeção* vem da idéia de projetar uma imagem em uma parede (com um projetor de *slides*, por exemplo). Imagine um feixe de luz com raios paralelos entre si e perpendiculares a **u** incidindo sobre **v**. A projeção de **v** sobre **u** é justamente a sombra lançada, ou projetada, por **v** sobre **u**.
- Pode ser útil interpretar $\text{proj}_{\mathbf{u}}(\mathbf{v})$ como uma função de variável **v**. Então, a variável **v** ocorre apenas uma vez no lado direito. Além disso, é de grande ajuda ter em mente a Figura 13, que nos lembra que $\text{proj}_{\mathbf{u}}(\mathbf{v})$ é um múltiplo escalar do vetor **u** (*não* **v**).
- Na definição de $\text{proj}_{\mathbf{u}}(\mathbf{v})$, embora tenhamos considerado **v** não nulo, assim como **u** (por quê?), é claro, do ponto de vista da geometria, que a projeção do vetor nulo sobre **u** é **0**. A definição confirma esse fato, pois $\left(\frac{\mathbf{u} \cdot \mathbf{0}}{\mathbf{u} \cdot \mathbf{u}}\right)\mathbf{u} = 0\mathbf{u} = \mathbf{0}$.
- Se o ângulo entre **u** e **v** for obtuso, como na Figura 14, então $\text{proj}_{\mathbf{u}}(\mathbf{v})$ terá sentido oposto ao de **u**, ou seja, $\text{proj}_{\mathbf{u}}(\mathbf{v})$ será um múltiplo escalar *negativo* de **u**.
- Se **u** for um vetor unitário, então $\text{proj}_{\mathbf{u}}(\mathbf{v}) = (\mathbf{u} \cdot \mathbf{v})\mathbf{u}$. (Por quê?)

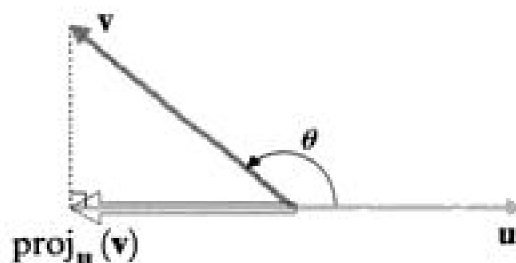


Figura 14

EXEMPLO 11 Ache a projeção de **v** sobre **u** em cada caso.

(a) $\mathbf{v} = \begin{bmatrix} -1 \\ 3 \end{bmatrix}$ e $\mathbf{u} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$.

(b) $\mathbf{v} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$ e $\mathbf{u} = \mathbf{e}_3$

(c) $\mathbf{v} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$ e $\mathbf{u} = \begin{bmatrix} 1/2 \\ 1/2 \\ 1/\sqrt{2} \end{bmatrix}$

SOLUÇÃO: (a) Calculamos $\mathbf{u} \cdot \mathbf{v} = \begin{bmatrix} 2 \\ 1 \end{bmatrix} \cdot \begin{bmatrix} -1 \\ 3 \end{bmatrix} = 1$ e $\mathbf{u} \cdot \mathbf{u} = \begin{bmatrix} 2 \\ 1 \end{bmatrix} \cdot \begin{bmatrix} 2 \\ 1 \end{bmatrix} = 5$. Portanto,

$$\text{proj}_{\mathbf{u}}(\mathbf{v}) = \left(\frac{\mathbf{u} \cdot \mathbf{v}}{\mathbf{u} \cdot \mathbf{u}}\right)\mathbf{u} = \frac{1}{5}\begin{bmatrix} 2 \\ 1 \end{bmatrix} = \begin{bmatrix} 2/5 \\ 1/5 \end{bmatrix}$$

Material com direitos autorais

(b) Como $\mathbf{e}_3$ é um vetor unitário,

$$\text{proj}_{\mathbf{e}_3}(\mathbf{v}) = (\mathbf{e}_3 \cdot \mathbf{v})\mathbf{e}_3 = 3\mathbf{e}_3 = \begin{bmatrix} 0 \\ 0 \\ 3 \end{bmatrix}$$

(c) Temos que $\|\mathbf{u}\| = \frac{1}{4} + \frac{1}{4} + \frac{1}{2} = 1$. Então,

$$\text{proj}_{\mathbf{u}}(\mathbf{v}) = (\mathbf{u} \cdot \mathbf{v})\mathbf{u} = \left(\frac{1}{2} + 1 + \frac{3}{\sqrt{2}}\right)\begin{bmatrix} 1/2 \\ 1/2 \\ 1/\sqrt{2} \end{bmatrix} = \frac{3(1+\sqrt{2})}{2}\begin{bmatrix} 1/2 \\ 1/2 \\ 1/\sqrt{2} \end{bmatrix} = \frac{3(1+\sqrt{2})}{4}\begin{bmatrix} 1 \\ 1 \\ \sqrt{2} \end{bmatrix}$$ ◆

◆ EXERCÍCIOS 1.3 ◆

Nos Exercícios de 1 a 6, calcule $\mathbf{u} \cdot \mathbf{v}$.

1. $\mathbf{u} = \begin{bmatrix} -1 \\ 2 \end{bmatrix}, \mathbf{v} = \begin{bmatrix} 3 \\ 1 \end{bmatrix}$ **2.** $\mathbf{u} = \begin{bmatrix} 3 \\ -2 \end{bmatrix}, \mathbf{v} = \begin{bmatrix} 4 \\ 6 \end{bmatrix}$

3. $\mathbf{u} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}, \mathbf{v} = \begin{bmatrix} 2 \\ 3 \\ 1 \end{bmatrix}$ **4.** $\mathbf{u} = \begin{bmatrix} 3{,}2 \\ -0{,}6 \\ -1{,}4 \end{bmatrix}, \mathbf{v} = \begin{bmatrix} 1{,}5 \\ 4{,}1 \\ -0{,}2 \end{bmatrix}$

5. $\mathbf{u} = [1, \sqrt{2}, \sqrt{3}, 0], \mathbf{v} = [4, -\sqrt{2}, 0, -5]$

6. $\mathbf{u} = [1{,}12, -3{,}25, 2{,}07, -1{,}83]$,
$\mathbf{v} = [-2{,}29, 1{,}72, 4{,}33, -1{,}54]$

Nos Exercícios de 7 a 12, determine a norma do vetor $\|\mathbf{u}\|$ *do exercício indicado e dê o versor de* $\mathbf{u}$.

7. Exercício 1 **8.** Exercício 2 **9.** Exercício 3

10. Exercício 4 **11.** Exercício 5 **12.** Exercício 6

Nos Exercícios de 13 a 16, encontre a distância $d(\mathbf{u}, \mathbf{v})$ *entre os vetores* $\mathbf{u}$ *e* $\mathbf{v}$ *do exercício indicado.*

13. Exercício 1 **14.** Exercício 2

15. Exercício 3 **16.** Exercício 4

17. Se $\mathbf{u}$, $\mathbf{v}$ e $\mathbf{w}$ são vetores em $\mathbb{R}^n$, $n \geq 2$ e c é um escalar, explique por que as seguintes expressões não fazem sentido:

(a) $\|\mathbf{u} \cdot \mathbf{v}\|$ (b) $\mathbf{u} \cdot \mathbf{v} + \mathbf{w}$
(c) $\mathbf{u} \cdot (\mathbf{v} \cdot \mathbf{w})$ (d) $c \cdot (\mathbf{v} + \mathbf{w})$

Nos Exercícios de 18 a 21, determine se o ângulo entre $\mathbf{u}$ *e* $\mathbf{v}$ *é agudo, obtuso ou reto.*

18. $\mathbf{u} = \begin{bmatrix} 3 \\ 0 \end{bmatrix}, \mathbf{v} = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$ **19.** $\mathbf{u} = \begin{bmatrix} 2 \\ -1 \\ 1 \end{bmatrix}, \mathbf{v} = \begin{bmatrix} 1 \\ -2 \\ -1 \end{bmatrix}$

20. $\mathbf{u} = [4, 3, -1], \mathbf{v} = [1, -1, 1]$

21. $\mathbf{u} = [0{,}9,\ 2{,}1,\ 1{,}2], \mathbf{v} = [-4{,}5,\ 2{,}6,\ -0{,}8]$

Nos Exercícios de 22 a 25, encontre o ângulo entre $\mathbf{u}$ *e* $\mathbf{v}$ *do exercício dado.*

22. Exercício 18 **23.** Exercício 19

24. Exercício 20 **25.** Exercício 21

26. Sejam $A = (-3, 2)$, $B = (1, 0)$ e $C = (4, 6)$. Prove que o triângulo ΔABC é um triângulo retângulo.

27. Sejam $A = (1, 1, -1)$, $B = (-3, 2, -2)$ e $C = (2, 2, -4)$. Prove que o triângulo ΔABC é um triângulo retângulo.

28. Determine o ângulo entre uma diagonal de um cubo e uma aresta adjacente.

29. Um cubo possui quatro diagonais. Mostre que não existem duas dessas diagonais que sejam perpendiculares entre si.

Nos Exercícios de 30 a 35, encontre a projeção de $\mathbf{v}$ *sobre* $\mathbf{u}$. *Faça um esboço nos Exercícios 30 e 31.*

30. $\mathbf{u} = \begin{bmatrix} 1 \\ -1 \end{bmatrix}, \mathbf{v} = \begin{bmatrix} 3 \\ -1 \end{bmatrix}$ **31.** $\mathbf{u} = \begin{bmatrix} \frac{3}{5} \\ -\frac{4}{5} \end{bmatrix}, \mathbf{v} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$

32. $\mathbf{u} = \begin{bmatrix} \frac{2}{3} \\ -\frac{2}{3} \\ -\frac{1}{3} \end{bmatrix}, \mathbf{v} = \begin{bmatrix} 2 \\ -2 \\ 2 \end{bmatrix}$ **33.** $\mathbf{u} = \begin{bmatrix} 1 \\ -1 \\ 1 \\ -1 \end{bmatrix}, \mathbf{v} = \begin{bmatrix} 2 \\ -3 \\ -1 \\ -2 \end{bmatrix}$

34. $\mathbf{u} = \begin{bmatrix} 0{,}5 \\ 1{,}5 \end{bmatrix}, \mathbf{v} = \begin{bmatrix} 2{,}1 \\ 1{,}2 \end{bmatrix}$

35. $\mathbf{u} = \begin{bmatrix} 3{,}01 \\ -0{,}33 \\ 2{,}52 \end{bmatrix}, \mathbf{v} = \begin{bmatrix} 1{,}34 \\ 4{,}25 \\ -1{,}66 \end{bmatrix}$

Material com direitos autorais

A Figura 15 sugere duas maneiras pelas quais os vetores podem ser usados para calcular a área de um triângulo. A área $\mathcal{A}$ do triângulo na parte (a) é dada por $\frac{1}{2}\|\mathbf{u}\|\|\mathbf{v} - \text{proj}_{\mathbf{u}}(\mathbf{v})\|$, e a parte (b) sugere a forma trigonométrica da área de um triângulo: $\mathcal{A} = \frac{1}{2}\|\mathbf{u}\|\|\mathbf{v}\|\,\text{sen}\,\theta$. (Podemos usar a identidade $\text{sen}\,\theta = \sqrt{1 - \cos^2\theta}$ para achar sen θ.) Nos Exercícios 36 e 37, calcule a área do triângulo de vértices dados, usando ambos os métodos.

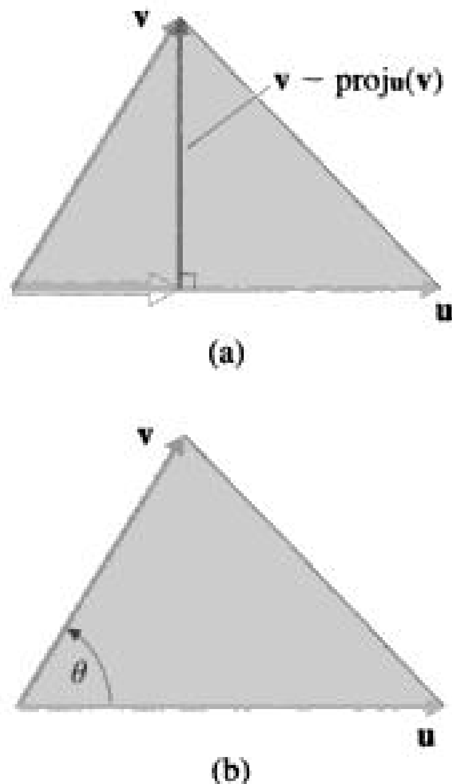


Figura 15

36. $A = (1, -1)$, $B = (2, 2)$, $C = (4, 0)$

37. $A = (3, -1, 4)$, $B = (4, -2, 6)$, $C = (5, 0, 2)$

Nos Exercícios 38 e 39, encontre todos os possíveis valores de k para os quais os dois vetores são ortogonais.

38. $\mathbf{u} = \begin{bmatrix} 2 \\ 3 \end{bmatrix}, \mathbf{v} = \begin{bmatrix} k+1 \\ k-1 \end{bmatrix}$ **39.** $\mathbf{u} = \begin{bmatrix} 1 \\ -1 \\ 2 \end{bmatrix}, \mathbf{v} = \begin{bmatrix} k^2 \\ k \\ -3 \end{bmatrix}$

40. Descreva todos os vetores $\mathbf{v} = \begin{bmatrix} x \\ y \end{bmatrix}$ que são ortogonais a $\mathbf{u} = \begin{bmatrix} 3 \\ 1 \end{bmatrix}$.

41. Descreva todos os vetores $\mathbf{v} = \begin{bmatrix} x \\ y \end{bmatrix}$ que são ortogonais a $\mathbf{u} = \begin{bmatrix} a \\ b \end{bmatrix}$.

42. Sob que condições as seguintes igualdades são verdadeiras para vetores **u** e **v** em $\mathbb{R}^2$ ou $\mathbb{R}^3$?

(a) $\|\mathbf{u} + \mathbf{v}\| = \|\mathbf{u}\| + \|\mathbf{v}\|$ (b) $\|\mathbf{u} + \mathbf{v}\| = \|\mathbf{u}\| - \|\mathbf{v}\|$

43. (a) Prove o Teorema 1(b). (b) Prove o Teorema 1(d).

44. Prove que $\mathbf{u} \cdot a\mathbf{v} = a(\mathbf{u} \cdot \mathbf{v})$, para todos os vetores **u** e **v** em $\mathbb{R}^n$ e todos os números a.

45. Prove que $\|\mathbf{u} - \mathbf{v}\| \geq \|\mathbf{u}\| - \|\mathbf{v}\|$, para quaisquer vetores **u** e **v** em $\mathbb{R}^n$. (*Sugestão*: substitua **u** por **u** − **v** na Desigualdade Triangular.)

46. Suponha conhecido que $\mathbf{u} \cdot \mathbf{v} = \mathbf{u} \cdot \mathbf{w}$. Pode-se concluir que $\mathbf{v} = \mathbf{w}$? Em caso afirmativo, dê uma prova válida em $\mathbb{R}^n$; caso contrário, dê um contra-exemplo (isto é, um conjunto específico de vetores **u**, **v** e **w** para os quais $\mathbf{u} \cdot \mathbf{v} = \mathbf{u} \cdot \mathbf{w}$ mas $\mathbf{v} \neq \mathbf{w}$).

47. Prove que $(\mathbf{u} + \mathbf{v}) \cdot (\mathbf{u} - \mathbf{v}) = \|\mathbf{u}\|^2 - \|\mathbf{v}\|^2$, para quaisquer vetores **u** e **v** em $\mathbb{R}^n$.

48. (a) Prove que $\|\mathbf{u} + \mathbf{v}\|^2 + \|\mathbf{u} - \mathbf{v}\|^2 = 2\|\mathbf{u}\|^2 + 2\|\mathbf{v}\|^2$, para quaisquer vetores **u** e **v** em $\mathbb{R}^n$.

(b) Desenhe um diagrama mostrando **u**, **v**, **u** + **v** e **u** − **v** em $\mathbb{R}^2$ e use a prova da parte (a) para deduzir uma propriedade sobre os paralelogramos.

49. Prove que $\mathbf{u} \cdot \mathbf{v} = \frac{1}{4}\|\mathbf{u} + \mathbf{v}\|^2 - \frac{1}{4}\|\mathbf{u} - \mathbf{v}\|^2$, para quaisquer vetores **u** e **v** em $\mathbb{R}^n$.

50. (a) Prove que $\|\mathbf{u} + \mathbf{v}\| = \|\mathbf{u} - \mathbf{v}\|$ se, e somente se, **u** e **v** são ortogonais.

(b) Desenhe um diagrama mostrando **u**, **v**, **u** + **v** e **u** − **v** em $\mathbb{R}^2$ e use a prova da parte (a) para deduzir uma propriedade sobre os paralelogramos.

51. (a) Prove que **u** + **v** e **u** − **v** são ortogonais se, e somente se, $\|\mathbf{u}\| = \|\mathbf{v}\|$. (*Sugestão*: veja o Exercício 47.)

(b) Desenhe um diagrama mostrando **u**, **v**, **u** + **v** e **u** − **v** em $\mathbb{R}^2$ e use a prova da parte (a) para deduzir uma propriedade sobre os paralelogramos.

52. (a) Prove que, se **u** é ortogonal a **v** e a **w**, então **u** é ortogonal a **v** + **w**.

(b) Prove que, se **u** é ortogonal a **v** e a **w**, então **u** é ortogonal a $s\mathbf{v} + t\mathbf{w}$, para quaisquer números s e t.

53. Prove que **u** é ortogonal a $\mathbf{v} - \text{proj}_{\mathbf{u}}(\mathbf{v})$, para quaisquer vetores **u** e **v** em $\mathbb{R}^n$, com $\mathbf{u} \neq \mathbf{0}$.

54. (a) Prove que $\text{proj}_{\mathbf{u}}(\text{proj}_{\mathbf{u}}(\mathbf{v})) = \text{proj}_{\mathbf{u}}(\mathbf{v})$.

(b) Prove que $\text{proj}_{\mathbf{u}}(\mathbf{v} - \text{proj}_{\mathbf{u}}(\mathbf{v})) = \mathbf{0}$.

(c) Explique geometricamente os resultados dos itens (a) e (b).

55. A Desigualdade de Cauchy-Schwarz $|\mathbf{u} \cdot \mathbf{v}| \leq \|\mathbf{u}\|\|\mathbf{v}\|$ é equivalente à desigualdade que obtemos ao elevarmos ao quadrado ambos os lados: $(\mathbf{u} \cdot \mathbf{v})^2 \leq \|\mathbf{u}\|^2\|\mathbf{v}\|^2$.

(a) Em $\mathbb{R}^2$, com $\mathbf{u} = \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}$ e $\mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}$, a desigualdade é escrita como

$$(u_1v_1 + u_2v_2)^2 \leq (u_1^2 + u_2^2)(v_1^2 + v_2^2)$$

Prove algebricamente essa desigualdade. (*Sugestão*: subtraia o lado esquerdo do lado direito e mostre que a diferença deve ser necessariamente não negativa.)

(b) Prove o análogo de (a) em $\mathbb{R}^3$.

Material com direitos autorais

56. Uma outra maneira de provar a Desigualdade de Cauchy-Schwarz é sugerida pela Figura 16, que mostra que, em $\mathbb{R}^2$ ou $\mathbb{R}^3$, $\|\text{proj}_{\mathbf{u}}(\mathbf{v})\| \leq \|\mathbf{v}\|$. Mostre que essa desigualdade é equivalente à Desigualdade de Cauchy-Schwarz.

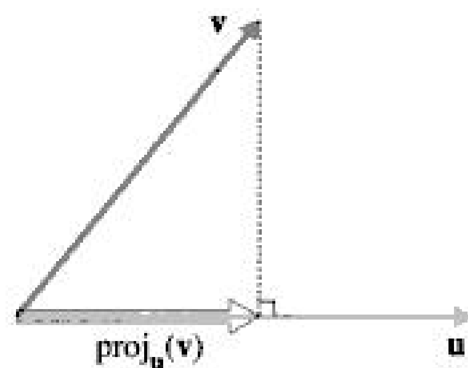


Figura 16

57. Use o fato de que $\text{proj}_{\mathbf{u}}(\mathbf{v}) = d\mathbf{u}$, para algum número d, juntamente com a Figura 17, para determinar d e, desse modo, deduzir a fórmula para $\text{proj}_{\mathbf{u}}(\mathbf{v})$.

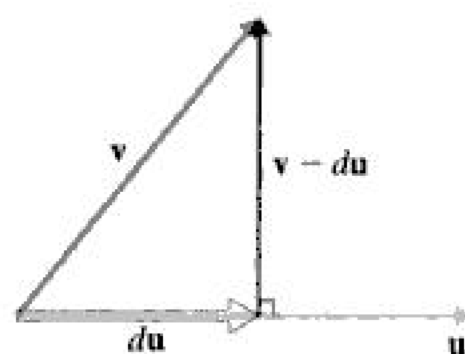


Figura 17

Material com direitos autorais

Investigação

Vetores e Geometria

Muitos resultados da geometria euclidiana plana podem ser demonstrados por meio de técnicas vetoriais. Por exemplo, no Exemplo 10 da *Seção 1.3*, usamos vetores para provar o Teorema de Pitágoras. Nesta investigação, usaremos vetores para provar alguns outros teoremas da geometria euclidiana.

Como introdução à notação e à linguagem vetorial para abordar a geometria, considere o seguinte exemplo fácil:

EXEMPLO 1 Dê uma descrição vetorial para o ponto médio M de um segmento de reta $\overline{AB}$.

SOLUÇÃO: Primeiramente, converteremos tudo para a notação vetorial. Se O denota a origem e P é um ponto, indicaremos por $\mathbf{p}$ o vetor $\overrightarrow{OP}$. Nessas condições, $\mathbf{a} = \overrightarrow{OA}$, $\mathbf{b} = \overrightarrow{OB}$, $\mathbf{m} = \overrightarrow{OM}$, e $\overrightarrow{AB} = \overrightarrow{OB} - \overrightarrow{OA} = \mathbf{b} - \mathbf{a}$ (Figura 1).

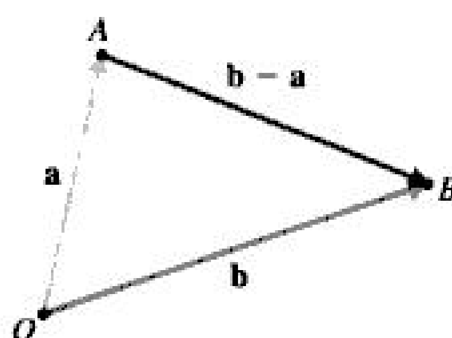


Figura 1

Como M é ponto médio de $\overline{AB}$, temos

$$\mathbf{m} - \mathbf{a} = \overrightarrow{AM} = \tfrac{1}{2}\overrightarrow{AB} = \tfrac{1}{2}(\mathbf{b} - \mathbf{a})$$

assim,

$$\mathbf{m} = \mathbf{a} + \tfrac{1}{2}(\mathbf{b} - \mathbf{a}) = \tfrac{1}{2}(\mathbf{a} + \mathbf{b})$$ ◆

1. Apresente uma descrição vetorial para o ponto P que está a um terço do percurso de A para B no segmento de reta $\overline{AB}$. Generalize.

2. Demonstre que o segmento de reta que une os pontos médios de dois lados de um triângulo é paralelo ao terceiro lado e tem a metade de seu comprimento. (Na notação vetorial, isso equivale a provar que $\overrightarrow{PQ} = \frac{1}{2}\overrightarrow{AB}$ na Figura 2.)

3. Demonstre que o quadrilátero $PQRS$ (Figura 3), cujos vértices são os pontos médios dos lados de um quadrilátero arbitrário $ABCD$, é um paralelogramo.

4. Uma ***mediana*** de um triângulo é um segmento de reta que une um vértice ao ponto médio do lado oposto (Figura 4). Prove que as três medianas de qualquer triângulo são *concorrentes* (isto é, têm um ponto comum de interseção) em um ponto G cuja distância a cada vértice é igual a dois terços da distância deste

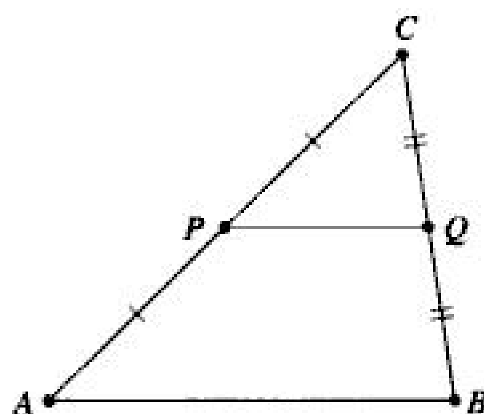


Figura 2

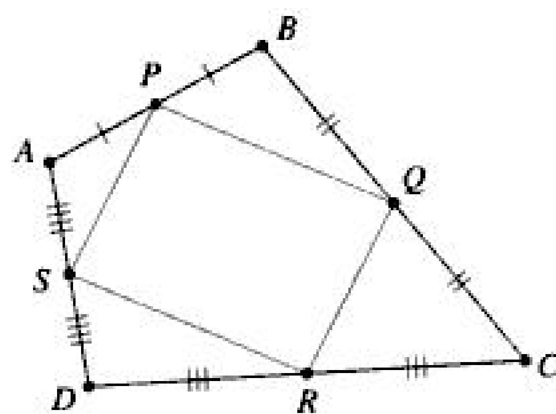


Figura 3

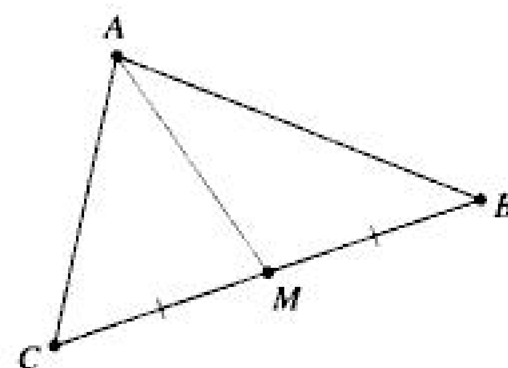


Figura 4 Uma mediana

Material com direitos autorais

vértice ao ponto médio do lado oposto. (*Sugestão*: na Figura 5, mostre que o ponto que está a dois terços da distância de A a P é dado por $\frac{1}{3}(\mathbf{a} + \mathbf{b} + \mathbf{c})$. Mostre então que o mesmo ponto está a dois terços da distância de B a Q e a dois terços da distância de C a R.) O ponto G da Figura 5 é chamado ***baricentro*** do triângulo.

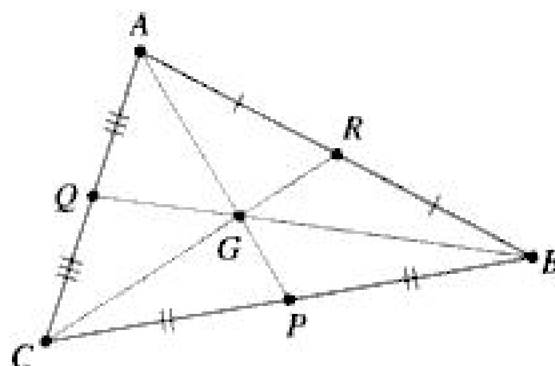


Figura 5

5. Uma ***altura*** de um triângulo é um segmento de reta que sai de um vértice e é perpendicular ao lado oposto (Figura 6). Prove que as três alturas de um triângulo são concorrentes em um mesmo ponto. (Sugestão: seja H o ponto de interseção das alturas correspondentes aos vértices A e B na Figura 7. Prove que $\overrightarrow{CH}$ é ortogonal a $\overrightarrow{AB}$.) O ponto H na Figura 7 é chamado ***ortocentro*** do triângulo.

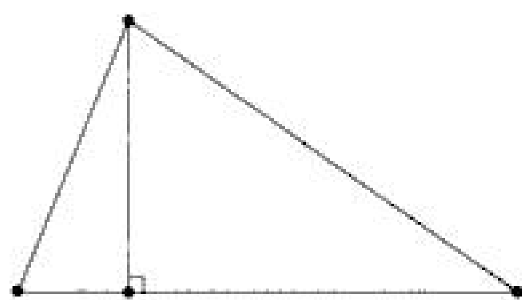


Figura 6 Uma altura

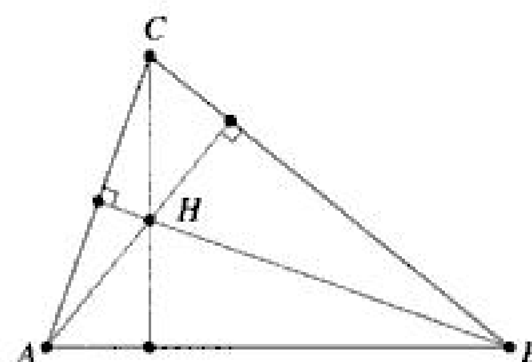


Figura 7

6. A ***mediatriz*** de um segmento de reta é a reta perpendicular a esse segmento e que passa por seu ponto médio (Figura 8). Prove que as mediatrizes dos três lados de um triângulo são concorrentes em um mesmo ponto. (Sugestão: seja K o ponto de interseção das mediatrizes de $\overline{AC}$ e $\overline{BC}$ na Figura 9. Prove que $\overrightarrow{RK}$ é ortogonal a $\overrightarrow{AB}$.) O ponto K na Figura 9 chama-se ***circuncentro*** do triângulo.

7. Sejam A e B as extremidades de um diâmetro de um círculo. Se C é um ponto qualquer do círculo, prove que $\angle ACB$ é um ângulo reto. (*Sugestão*: na Figura 10, seja O o centro do círculo. Expresse todos os elementos em termos de $\mathbf{a}$ e $\mathbf{c}$ e mostre que $\overrightarrow{AC}$ é ortogonal a $\overrightarrow{BC}$.)

8. Prove que os segmentos de reta que unem os pontos médios dos lados opostos de um quadrilátero se interceptam de modo a dividir cada um deles em duas partes de mesmo comprimento (Figura 11).

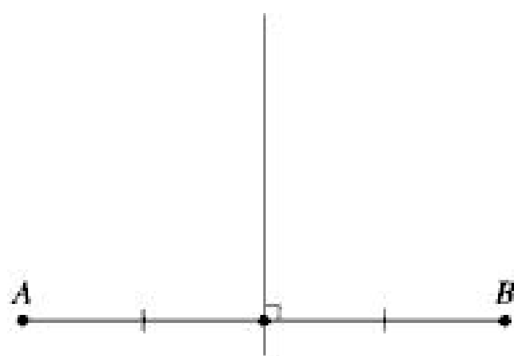


Figura 8 Uma mediatriz

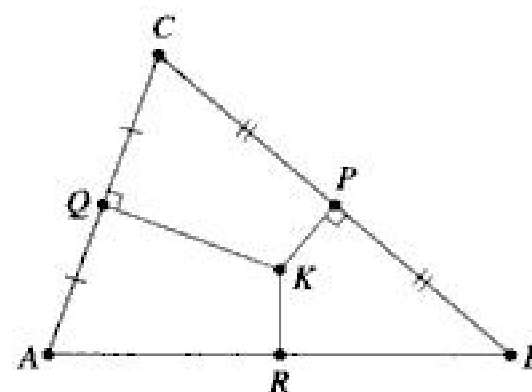


Figura 9

Material com direitos autorais

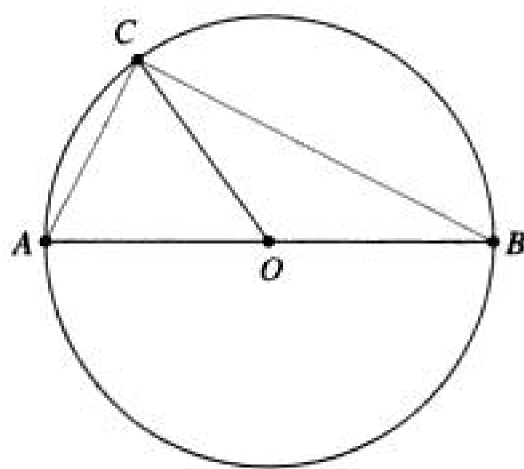


Figura 10

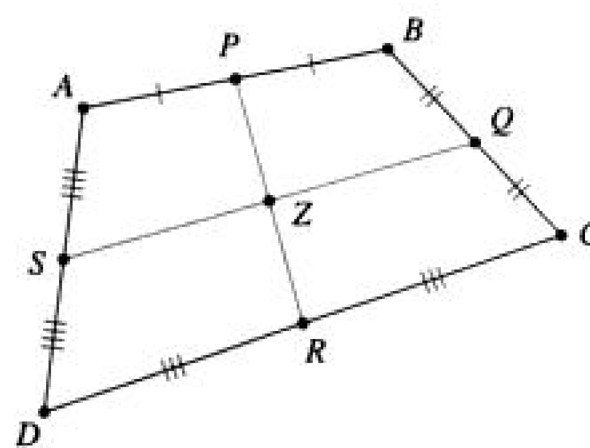


Figura 11

1.4 Retas e Planos

Todos estamos familiarizados com a equação de uma reta no plano cartesiano. Agora, vamos considerar retas em $\mathbb{R}^2$ do ponto de vista vetorial. As idéias que aparecem neste estudo nos permitem generalizar para retas em $\mathbb{R}^3$ e depois para planos em $\mathbb{R}^3$. Muito da álgebra linear que estudaremos nos capítulos posteriores tem sua origem na geometria simples de retas e planos; a habilidade de visualizar esses elementos e pensar em um problema do ponto de vista geométrico será de grande valia.

Retas em $\mathbb{R}^2$ e em $\mathbb{R}^2$

No plano xy, a forma geral da equação de uma reta é $ax + by = c$. Se $b \neq 0$, a equação pode ser reescrita como $y = -(a/b)x + c/b$, que tem a forma $y = mx + k$. [Esta é a equação da reta na forma reduzida: m representa a inclinação da reta, e o ponto de coordenadas $(0, k)$ é sua interseção com o eixo y.] Para fazer uma descrição com vetores, vamos considerar um exemplo.

EXEMPLO 1 A reta r de equação $2x + y = 0$ é mostrada na Figura 1. É uma reta de inclinação -2 passando pela origem. O lado esquerdo da equação tem a forma de um produto escalar; de fato, se consideramos $\mathbf{n} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$ e $\mathbf{x} = \begin{bmatrix} x \\ y \end{bmatrix}$, a equação se torna $\mathbf{n} \cdot \mathbf{x} = 0 \cdot \mathbf{x} = 0$. O vetor $\mathbf{n}$ é ortogonal à reta – isto é, *ortogonal* a qual-

A palavra latina *norma* se refere ao esquadro de carpinteiro, usado para desenhar ângulos retos. Assim, um vetor *normal* é um vetor que é perpendicular a algum outro objeto, geralmente um plano.

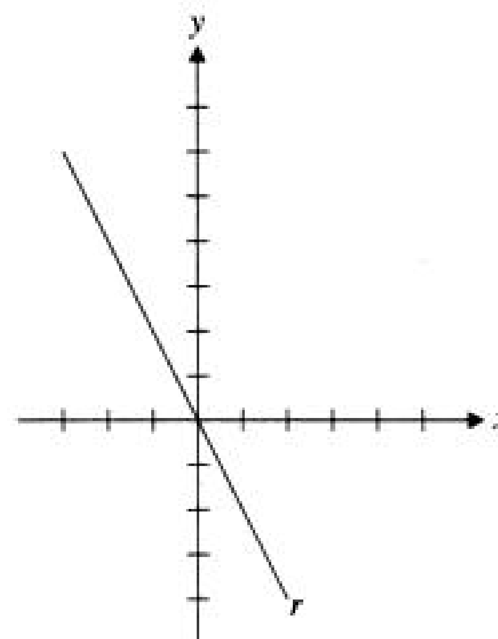


Figura 1 $2x + y = 0$

Material com direitos autorais

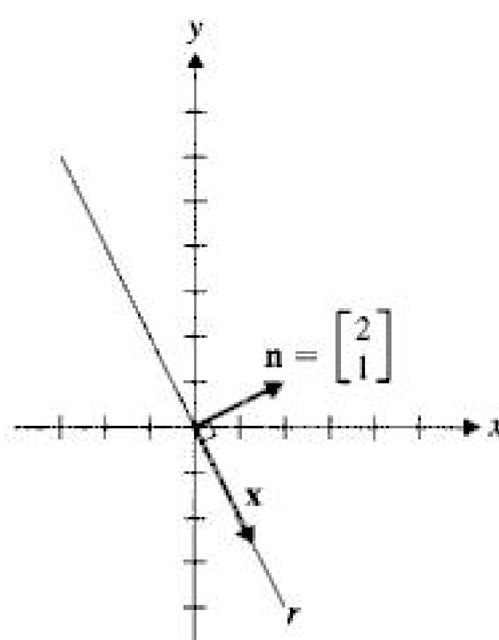


Figura 2 O vetor normal **n**

quer vetor **x** paralelo à reta (Figura 2) – e é chamado ***vetor normal*** à reta. A equação $\mathbf{n} \cdot \mathbf{x} = 0$ é a equação da reta r na *forma normal*. ◆

Uma outra maneira de pensar sobre essa reta é imaginar uma partícula movendo-se ao longo dela. Suponhamos que a partícula esteja na origem no instante $t = 0$ e se movimente ao longo da reta de modo que sua coordenada x varie em uma unidade por segundo. Então, no instante $t = 1$, a partícula está em $(1, -2)$; em

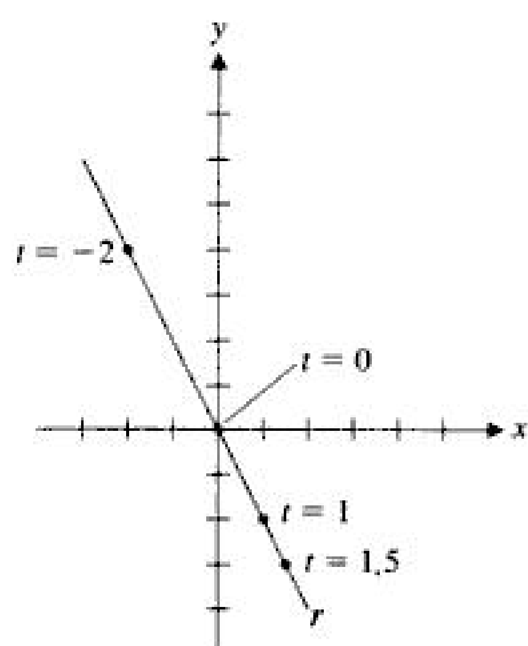


Figura 3

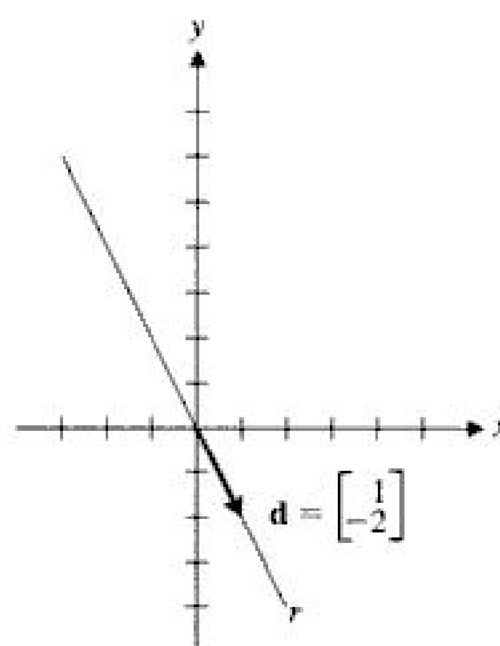


Figura 4 O vetor diretor **d**

$t = 1{,}5$, está em $(1{,}5, -3)$; se permitirmos valores negativos para t (isto é, se levarmos em conta onde a partícula estava no passado), em $t = -2$ estará (ou esteve) em $(-2, 4)$. Tal movimento está ilustrado na Figura 3. Em geral, se $x = t$, então $y = -2t$, e podemos escrever essa relação na forma vetorial como

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} t \\ -2t \end{bmatrix} = t\begin{bmatrix} 1 \\ -2 \end{bmatrix}$$

Qual é o significado do vetor $\mathbf{d} = \begin{bmatrix} 1 \\ -2 \end{bmatrix}$? Trata-se de um particular vetor paralelo a r, chamado ***vetor diretor*** da reta. Como mostra a Figura 4, podemos escrever a equação de r como $\mathbf{x} = t\mathbf{d}$. Essa é uma *equação vetorial* da reta.

Se a reta não passar pela origem, temos de fazer algumas pequenas modificações.

Material com direitos autorais

EXEMPLO 2 Considere a reta r de equação $2x + y = 5$ (Figura 5). Trata-se apenas de uma translação em cinco unidades para cima, da reta do Exemplo 1. Ela também possui inclinação -2, mas sua interseção com o eixo y é o ponto $(0, 5)$. É claro que os vetores $\mathbf{d}$ e $\mathbf{n}$ do Exemplo 1 são, respectivamente, um vetor diretor e um vetor normal para essa reta também.

Assim, $\mathbf{n}$ é ortogonal a todo vetor que seja paralelo a r. O ponto $P = (1, 3)$ pertence a r. Se $X = (x, y)$ representa um ponto genérico de r, o vetor $\overrightarrow{PX} = \mathbf{x} - \mathbf{p}$ é paralelo a r e $\mathbf{n} \cdot (\mathbf{x} - \mathbf{p}) = 0$ (veja a Figura 6). Simplificando, temos $\mathbf{n} \cdot \mathbf{x} = \mathbf{n} \cdot \mathbf{p}$. Para verificar, calculamos

$$\mathbf{n} \cdot \mathbf{x} = \begin{bmatrix} 2 \\ 1 \end{bmatrix} \cdot \begin{bmatrix} x \\ y \end{bmatrix} = 2x + y \quad \text{e} \quad \mathbf{n} \cdot \mathbf{p} = \begin{bmatrix} 2 \\ 1 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 3 \end{bmatrix} = 5$$

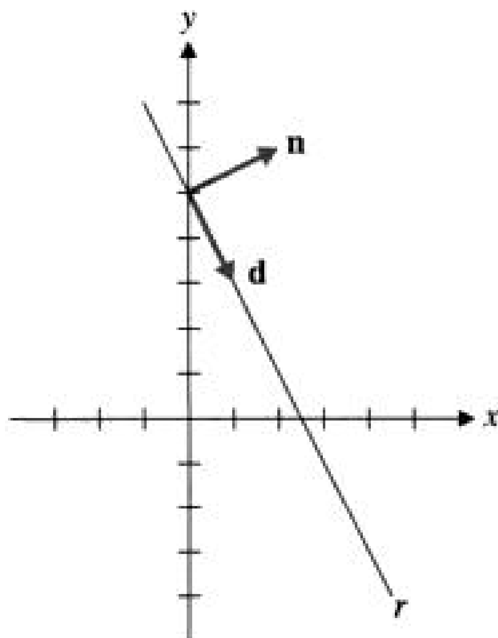


Figura 5 $2x + y = 5$

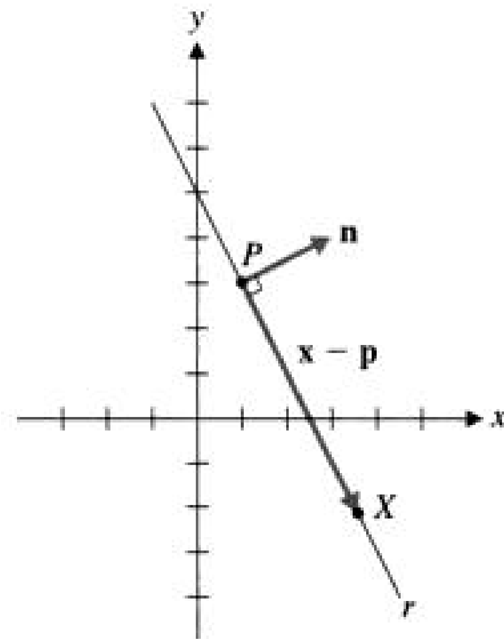


Figura 6 $\mathbf{n} \cdot (\mathbf{x} - \mathbf{p}) = 0$

Assim, a equação normal $\mathbf{n} \cdot \mathbf{x} = \mathbf{n} \cdot \mathbf{p}$ é uma representação um pouco diferente da equação da reta na forma geral. (Note que, no Exemplo 1, $\mathbf{p}$ era o vetor nulo, e o lado direito da equação era $\mathbf{n} \cdot \mathbf{p} = 0$.) ◆

Esses resultados conduzem à seguinte definição:

Definição Dada uma reta r em $\mathbb{R}^2$, uma equação escrita na forma

$$\mathbf{n} \cdot (\mathbf{x} - \mathbf{p}) = 0 \quad \text{ou} \quad \mathbf{n} \cdot \mathbf{x} = \mathbf{n} \cdot \mathbf{p}$$

onde $\mathbf{p}$ é um ponto específico de r e $\mathbf{n} \neq \mathbf{0}$ é um vetor normal a r é chamada **equação da reta r na forma normal.**

Uma equação escrita na forma $ax + by = c$, onde $\mathbf{n} = \begin{bmatrix} a \\ b \end{bmatrix}$ é um vetor normal a r, é chamada **equação da reta r na forma geral.**

Continuando com o Exemplo 2, vamos encontrar uma equação de r na forma vetorial. Observe que, para cada escolha de $\mathbf{x}$, $\mathbf{x} - \mathbf{p}$ deve ser paralelo ao vetor diretor $\mathbf{d}$, e, portanto, um múltiplo de $\mathbf{d}$. Isto é, $\mathbf{x} - \mathbf{p} = t\mathbf{d}$ ou $\mathbf{x} = \mathbf{p} + t\mathbf{d}$ para algum escalar t. Em coordenadas temos

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 1 \\ 3 \end{bmatrix} + t\begin{bmatrix} 1 \\ -2 \end{bmatrix} \tag{1}$$

ou

$$\begin{aligned} x &= 1 + t \\ y &= 3 - 2t \end{aligned} \tag{2}$$

A equação (1) é a forma vetorial da equação de r, e as equações (2), das coordenadas, são chamadas *equações paramétricas* da reta. A variável t é chamada *parâmetro*.

Material com direitos autorais

A palavra *parâmetro* e o adjetivo correspondente paramétrico vêm das palavras gregas *para*, que significa "lado a lado", e *metron*, que quer dizer "medida". Matematicamente falando, um parâmetro é uma variável em termos da qual se expressam outras variáveis – uma nova "medida" colocada ao lado de outras antigas.

Como generalizar para $\mathbb{R}^3$? Observe que as equações de reta nas formas vetorial e paramétrica se estendem perfeitamente. A noção de inclinação de uma reta em $\mathbb{R}^2$ – que é difícil de generalizar para três dimensões – é substituída pela noção mais conveniente de vetor diretor, conduzindo à seguinte definição:

Definição Em $\mathbb{R}^2$ ou $\mathbb{R}^3$, se $\mathbf{p}$ for um ponto de r e $\mathbf{d} \neq \mathbf{0}$ for um vetor diretor de r, a equação

$$\mathbf{x} = \mathbf{p} + t\mathbf{d}$$

é chamada ***equação vetorial da reta r***.

As equações correspondentes às coordenadas são chamadas ***equações paramétricas*** de r.

Freqüentemente, abreviaremos um pouco essa terminologia, referindo-nos simplesmente à equação geral, normal, vetorial e às equações paramétricas de uma reta ou plano.

EXEMPLO 3 Determine equações vetorial e paramétricas para a reta em $\mathbb{R}^3$ que passa pelo ponto $P = (1, 2, -1)$ e é paralela ao vetor $\mathbf{d} = \begin{bmatrix} 5 \\ -1 \\ 3 \end{bmatrix}$.

SOLUÇÃO: A equação vetorial $\mathbf{x} = \mathbf{p} + t\mathbf{d}$ pode ser escrita como

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix} + t\begin{bmatrix} 5 \\ -1 \\ 3 \end{bmatrix}$$

As equações paramétricas são

$$\begin{aligned} x &= 1 + 5t \\ y &= 2 - t \\ z &= -1 + 3t \end{aligned}$$

◆

Observações:

- A equação vetorial e as equações paramétricas de uma dada reta r não são únicas – de fato, existem infinitas, já que podemos considerar qualquer ponto de r como $\mathbf{p}$ e qualquer vetor diretor de r. No entanto, todos os vetores diretores são claramente múltiplos uns dos outros.

No Exemplo 3, $(6, 1, 2)$ é um outro ponto sobre a reta (considere $t = 1$), e $\begin{bmatrix} 10 \\ -2 \\ 6 \end{bmatrix}$ é outro vetor diretor. Dessa maneira,

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 6 \\ 1 \\ 2 \end{bmatrix} + s\begin{bmatrix} 10 \\ -2 \\ 6 \end{bmatrix}$$

é uma equação vetorial diferente (mas equivalente) para a reta. A relação entre os dois parâmetros s e t pode ser obtida comparando-se as equações paramétricas: para cada ponto (x, y, z) de r, temos

$$\begin{aligned} x &= 1 + 5t = 6 + 10s \\ y &= 2 - t = 1 - 2s \\ z &= -1 + 3t = 2 + 6s \end{aligned}$$

Material com direitos autorais

implicando que

$$\begin{aligned} -10s + 5t &= 5 \\ 2s - t &= -1 \\ -6s + 3t &= 3 \end{aligned}$$

Todas essas equações se reduzem a $t = 1 + 2s$.

◆ Intuitivamente, sabemos que uma reta é um objeto *unidimensional*. A idéia de "dimensão" será explicada nos Capítulos 3 e 6, mas, para o momento, observe que essa idéia parece combinar com o fato de que a equação vetorial de uma reta requer apenas *um* parâmetro.

EXEMPLO 4 É freqüente a expressão "um par de pontos determina uma reta". Encontre uma equação vetorial para a reta r em $\mathbb{R}^3$ determinada pelos pontos $P = (-1, 5, 0)$ e $Q = (2, 1, 1)$.

SOLUÇÃO: Podemos escolher qualquer ponto de r como $\mathbf{p}$; então, usaremos P (Q também seria bom). Um vetor diretor conveniente é $\mathbf{d} = \overrightarrow{PQ} = \begin{bmatrix} 3 \\ -4 \\ 1 \end{bmatrix}$ (ou qualquer múltiplo escalar deste). Assim, obtemos

$$\begin{aligned} \mathbf{x} &= \mathbf{p} + t\mathbf{d} \\ &= \begin{bmatrix} -1 \\ 5 \\ 0 \end{bmatrix} + t\begin{bmatrix} 3 \\ -4 \\ 1 \end{bmatrix} \end{aligned}$$ ◆

Planos em $\mathbb{R}^3$

A próxima pergunta que devemos fazer é "Como a equação geral de uma reta se generaliza para $\mathbb{R}^3$?". É razoável supor que, se $ax + by = c$ é a forma geral da equação de uma reta em $\mathbb{R}^2$, $ax + by + cz = d$ deva representar uma reta em $\mathbb{R}^3$. Na forma normal, essa equação é $\mathbf{n} \cdot \mathbf{x} = \mathbf{n} \cdot \mathbf{p}$, onde $\mathbf{n}$ é um vetor normal à reta e $\mathbf{p}$ corresponde a um ponto sobre a reta.

Para verificar se essa é uma hipótese razoável, vamos refletir sobre o caso especial da equação $ax + by + cz = 0$. Na forma normal, ela é $\mathbf{n} \cdot \mathbf{x} = 0$, onde $\mathbf{n} = \begin{bmatrix} a \\ b \\ c \end{bmatrix}$.

No entanto, o conjunto de todos os vetores $\mathbf{x}$ que satisfazem essa equação é o conjunto de todos os vetores ortogonais a $\mathbf{n}$. Como mostra a Figura 7, infinitos vetores de direções diferentes possuem essa propriedade, determinando uma família de *planos* paralelos. Dessa forma, nossa suposição está incorreta: parece que $ax + by + cz = d$ é a equação de um plano – não de uma reta – em $\mathbb{R}^3$.

Tornemos mais precisa essa descoberta. Todo plano $\mathcal{P}$ em $\mathbb{R}^3$ pode ser determinado especificando-se um ponto $\mathbf{p}$ em $\mathcal{P}$ e um vetor não nulo $\mathbf{n}$ normal a $\mathcal{P}$ (Figura 8). Dessa maneira, se $\mathbf{x}$ representa um ponto arbitrário de $\mathcal{P}$, temos que $\mathbf{n} \cdot (\mathbf{x} - \mathbf{p}) = 0$ ou $\mathbf{n} \cdot \mathbf{x} = \mathbf{n} \cdot \mathbf{p}$. Se $\mathbf{n} = \begin{bmatrix} a \\ b \\ c \end{bmatrix}$ e $\mathbf{x} = \begin{bmatrix} x \\ y \\ z \end{bmatrix}$, então, em termos das coordenadas, a equação se torna $ax + by + cz = d$ (onde $d = \mathbf{n} \cdot \mathbf{p}$).

Material com direitos autorais

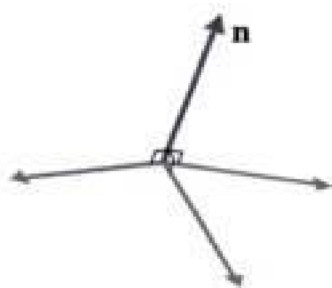


Figura 7 **n** é ortogonal a infinitos vetores

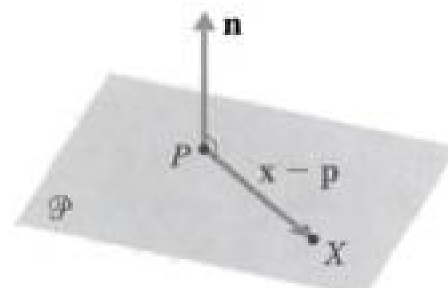


Figura 8 $\mathbf{n}\cdot(\mathbf{x}-\mathbf{p})=0$

Definição Seja 𝒫 um plano em $\mathbb{R}^3$, p um ponto pertencente a 𝒫 e $\mathbf{n}\neq 0$ um **vetor normal** a 𝒫. Uma equação escrita na forma

$$\mathbf{n}\cdot(\mathbf{x}-\mathbf{p})=0 \quad \text{ou} \quad \mathbf{n}\cdot\mathbf{x}=\mathbf{n}\cdot\mathbf{p}$$

é chamada **equação normal de** 𝒫. Uma equação escrita na forma $ax+by+cz=d$, onde $\mathbf{n}=\begin{bmatrix} a\\ b\\ c\end{bmatrix}$, é chamada **equação geral de** 𝒫.

Note que qualquer múltiplo escalar de um vetor normal a um plano é um outro vetor normal.

EXEMPLO 5 Ache equações nas formas normal e geral para o plano que contém o ponto $P=(6,0,1)$ e tem vetor normal $\mathbf{n}=\begin{bmatrix}1\\2\\3\end{bmatrix}$.

SOLUÇÃO: Com $\mathbf{p}=\begin{bmatrix}6\\0\\1\end{bmatrix}$ e $\mathbf{x}=\begin{bmatrix}x\\y\\z\end{bmatrix}$, temos $\mathbf{n}\cdot\mathbf{p}=1\cdot 6+2\cdot 0+3\cdot 1=9$; portanto, a equação normal $\mathbf{n}\cdot\mathbf{x}=\mathbf{n}\cdot\mathbf{p}$ se transforma na equação geral $x+2y+3z=9$. ◆

Geometricamente, é claro que planos paralelos têm os mesmos vetores normais. Assim, suas equações gerais têm os lados esquerdos múltiplos uns dos outros. Por exemplo, $2x+4y+6z=10$ é uma equação geral de um plano paralelo ao plano do Exemplo 5, já que podemos reescrever essa equação como $x+2y+3z=5$ – de onde percebemos que os dois planos têm o mesmo vetor normal **n**. (Note que os planos não coincidem, pois os lados direitos das equações dos dois planos são distintos.)

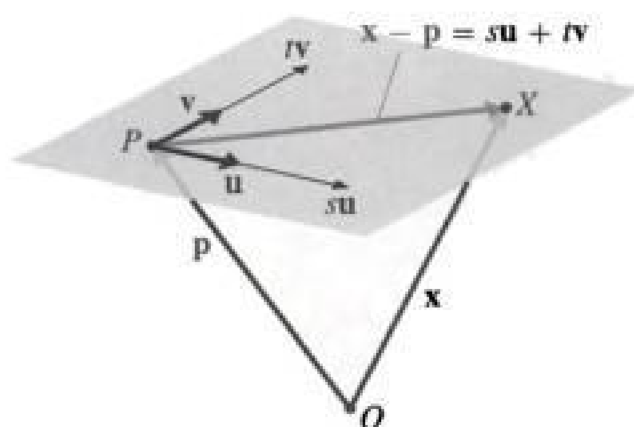


Figura 9 $\mathbf{x}-\mathbf{p}=s\mathbf{u}+t\mathbf{v}$

Material com direitos autorais

Podemos também expressar a equação do plano na forma vetorial ou na forma paramétrica. Para isso, observe que um plano pode ser determinado especificando-se um de seus pontos P (pelo vetor $\mathbf{p}$) e *dois* vetores diretores $\mathbf{u}$ e $\mathbf{v}$ paralelos ao plano (mas não paralelos entre si). Como mostra a Figura 9, dado qualquer ponto X do plano (localizado pelo vetor $\mathbf{x}$), podemos sempre encontrar múltiplos apropriados $s\mathbf{u}$ e $t\mathbf{v}$ dos vetores diretores, de maneira que $\mathbf{x} - \mathbf{p} = s\mathbf{u} + t\mathbf{v}$ ou $\mathbf{x} = \mathbf{p} + s\mathbf{u} + t\mathbf{v}$. Se escrevermos essa equação em coordenadas, obteremos equações paramétricas para o plano.

Definição Uma ***equação vetorial para o plano*** $\mathcal{P}$ em $\mathbb{R}^3$ é uma equação que pode ser escrita na forma

$$\mathbf{x} = \mathbf{p} + s\mathbf{u} + t\mathbf{v}$$

onde $\mathbf{p}$ é um ponto de $\mathcal{P}$ e $\mathbf{u}$ e $\mathbf{v}$ são vetores diretores de $\mathcal{P}$ ($\mathbf{u}$ e $\mathbf{v}$ são não nulos e paralelos a $\mathcal{P}$, mas não paralelos entre si).

As equações correspondentes às coordenadas na equação vetorial são chamadas ***equações paramétricas*** de $\mathcal{P}$.

EXEMPLO 6 Encontre uma equação vetorial e equações paramétricas para o plano do Exemplo 5.

SOLUÇÃO: Precisamos determinar dois vetores diretores. Temos um ponto $P = (6, 0, 1)$ no plano; se encontrarmos dois outros pontos Q e R em $\mathcal{P}$, os vetores $\overrightarrow{PQ}$ e $\overrightarrow{PR}$ poderão servir como vetores diretores (a menos que, por falta de sorte, eles sejam paralelos!). Por tentativa e erro, observe que os pontos $Q = (9, 0, 0)$ e $R = (3, 3, 0)$ satisfazem a equação geral $x + 2y + 3z = 9$, e, portanto, estão no plano. Calculamos

$$\mathbf{u} = \overrightarrow{PQ} = \mathbf{q} - \mathbf{p} = \begin{bmatrix} 3 \\ 0 \\ -1 \end{bmatrix} \quad \text{e} \quad \mathbf{v} = \overrightarrow{PR} = \mathbf{r} - \mathbf{p} = \begin{bmatrix} -3 \\ 3 \\ -1 \end{bmatrix}$$

que, por não serem múltiplos escalares um do outro, servirão como vetores diretores. Por conseguinte, obtemos uma equação vetorial de $\mathcal{P}$,

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 6 \\ 0 \\ 1 \end{bmatrix} + s\begin{bmatrix} 3 \\ 0 \\ -1 \end{bmatrix} + t\begin{bmatrix} -3 \\ 3 \\ -1 \end{bmatrix}$$

e as correspondentes equações paramétricas,

$$\begin{aligned} x &= 6 + 3s - 3t \\ y &= \qquad\quad\; 3t \\ z &= 1 - \;\, s - \;\, t \end{aligned}$$

[O que aconteceria se escolhêssemos $R = (0, 0, 3)$?] ◆

Observações:

- Um plano é um objeto bidimensional, e suas equações, tanto vetoriais quanto paramétricas, requerem *dois* parâmetros.

Material com direitos autorais

- Como mostra a Figura 7, dado um ponto P e um vetor não nulo $\mathbf{n}$ em $\mathbb{R}^3$, existem infinitas retas que passam por P tendo $\mathbf{n}$ como vetor normal. Por outro lado, P e dois vetores normais não paralelos, $\mathbf{n}_1$ e $\mathbf{n}_2$, servem para determinar uma reta r univocamente, pois r é a reta que passa por P e é perpendicular ao plano de equação $\mathbf{x} = \mathbf{p} + s\mathbf{n}_1 + t\mathbf{n}_2$ (Figura 10). Dessa maneira, uma reta em $\mathbb{R}^3$ pode ser identificada por um par de equações

$$a_1x + b_1y + c_1z = d_1$$
$$a_2x + b_2y + c_2z = d_2$$



cada uma delas correspondendo a um vetor normal. Como tais equações identificam um par de planos não paralelos (por que não paralelos?), essa é a descrição de uma reta como interseção de dois planos não paralelos (Figura 11). Algebricamente, a reta consiste em todos os pontos (x, y, z) que satisfazem simultaneamente ambas as equações. Exploraremos mais esse conceito no Capítulo 2, quando discutirmos a solução de um sistema de equações lineares.

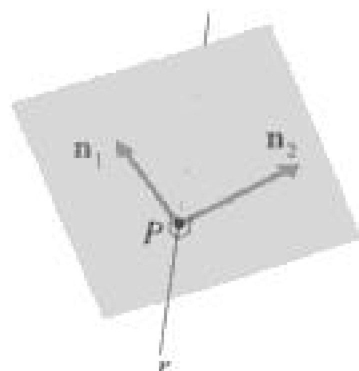


Figura 10 Um ponto e dois vetores normais determinam uma reta r

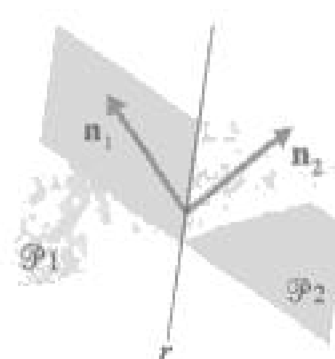


Figura 11 Uma reta interseção de dois planos

As Tabelas 1 e 2 resumem as informações apresentadas até agora sobre equações de retas e de planos.

Tabela 1 Equações de Retas em $\mathbb{R}^2$

Forma Normal	Forma Geral	Forma Vetorial	Forma Paramétrica
$\mathbf{n} \cdot \mathbf{x} = \mathbf{n} \cdot \mathbf{p}$	$ax + by = c$	$\mathbf{x} = \mathbf{p} + t\mathbf{d}$	$x = p_1 + td_1$ $y = p_2 + td_2$

Tabela 2 Retas e Planos em $\mathbb{R}^3$

	Forma Normal	Forma Geral	Forma Vetorial	Forma Paramétrica
Retas	$\mathbf{n}_1 \cdot \mathbf{x} = \mathbf{n}_1 \cdot \mathbf{p}_1$ $\mathbf{n}_2 \cdot \mathbf{x} = \mathbf{n}_2 \cdot \mathbf{p}_2$	$a_1x + b_1y + c_1z = d_1$ $a_2x + b_2y + c_2z = d_2$	$\mathbf{x} = \mathbf{p} + t\mathbf{d}$	$x = p_1 + td_1$ $y = p_2 + td_2$ $z = p_3 + td_3$
Planos	$\mathbf{n} \cdot \mathbf{x} = \mathbf{n} \cdot \mathbf{p}$	$ax + by + cz = d$	$\mathbf{x} = \mathbf{p} + s\mathbf{u} + t\mathbf{v}$	$x = p_1 + su_1 + tv_1$ $y = p_2 + su_2 + tv_2$ $z = p_3 + su_3 + tv_3$

Material com direitos autorais

Observe mais uma vez que uma equação geral em $\mathbb{R}^2$ descreve uma reta, mas em $\mathbb{R}^3$ descreve um plano. [Em dimensões maiores, um objeto (reta, plano etc.) determinado por uma única equação desse tipo é usualmente chamado ***hiperplano***.] A relação entre as dimensões do objeto, o número de equações necessárias para descrevê-lo e a dimensão do espaço é descrita pela seguinte "fórmula de equilíbrio":

(*dimensão* do objeto) + (*número* de equações na forma geral) = dimensão do espaço

Quanto maior a dimensão do objeto, menor o número de equações necessárias. Por exemplo, um plano em $\mathbb{R}^3$ é bidimensional, requer uma equação geral e está no espaço tridimensional: $2 + 1 = 3$. Uma reta em $\mathbb{R}^3$ é unidimensional e, assim, precisa de $3 - 1 = 2$ equações para ser identificada. Note que a dimensão do objeto também combina com a quantidade de parâmetros da equação na forma vetorial ou paramétrica. Noções de "dimensão" serão estudadas nos Capítulos 3 e 6, mas, por enquanto, essas observações intuitivas nos servirão.

Vamos agora achar a distância de um ponto a uma reta ou um plano relacionando os resultados da Seção 1.3 com os desta seção.

EXEMPLO 7 Determine a distância do ponto $B = (1, 0, 2)$ à reta r que passa pelo ponto $A = (3, 1, 1)$ e tem vetor diretor $\mathbf{d} = \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}$.

SOLUÇÃO: Como já sabemos, é necessário calcular o comprimento de $\overrightarrow{PB}$, onde P é o ponto de r no pé da perpendicular a partir de B. Se designarmos $\mathbf{v} = \overrightarrow{AB}$, então $\overrightarrow{AP} = \text{proj}_{\mathbf{d}}(\mathbf{v})$ e $\overrightarrow{PB} = \mathbf{v} - \text{proj}_{\mathbf{d}}(\mathbf{v})$ (veja a Figura 12). Faremos os cálculos necessários em várias etapas.

Etapa 1.
$$\mathbf{v} = \overrightarrow{AB} = \mathbf{b} - \mathbf{a} = \begin{bmatrix} 1 \\ 0 \\ 2 \end{bmatrix} - \begin{bmatrix} 3 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} -2 \\ -1 \\ 1 \end{bmatrix}$$

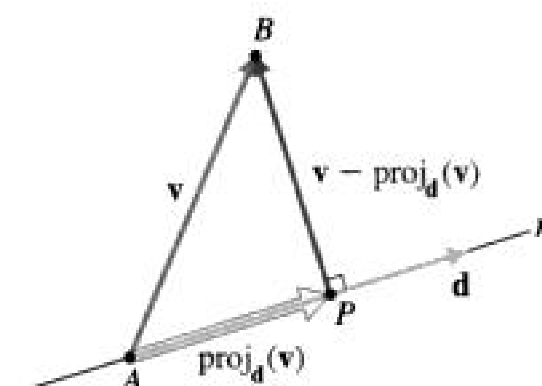


Figura 12 $d(B, r) = \|\mathbf{v} - \text{proj}_{\mathbf{d}}(\mathbf{v})\|$

Etapa 2. A projeção de $\mathbf{v}$ sobre $\mathbf{d}$ é

$$\begin{aligned}
\text{proj}_{\mathbf{d}}(\mathbf{v}) &= \left(\frac{\mathbf{d}\cdot\mathbf{v}}{\mathbf{d}\cdot\mathbf{d}}\right)\mathbf{d} \\
&= \left(\frac{(-1)\cdot(-2) + 1\cdot(-1) + 0\cdot 1}{(-1)^2 + 1 + 0}\right)\begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} \\
&= \tfrac{1}{2}\begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} \\
&= \begin{bmatrix} -\frac{1}{2} \\ \frac{1}{2} \\ 0 \end{bmatrix}
\end{aligned}$$

Material com direitos autorais

Etapa 3. O vetor que queremos é

$$\mathbf{v} - \text{proj}_{\mathbf{d}}(\mathbf{v}) = \begin{bmatrix} -2 \\ -1 \\ 1 \end{bmatrix} - \begin{bmatrix} -\frac{1}{2} \\ \frac{1}{2} \\ 0 \end{bmatrix} = \begin{bmatrix} -\frac{3}{2} \\ -\frac{3}{2} \\ 1 \end{bmatrix}$$

Etapa 4. A distância d(B, r) de B a r é

$$\|\mathbf{v} - \text{proj}_{\mathbf{d}}(\mathbf{v})\| = \left\| \begin{bmatrix} -\frac{3}{2} \\ -\frac{3}{2} \\ 1 \end{bmatrix} \right\|$$

Usando o Teorema 2(b) para simplificar os cálculos, temos

$$\begin{aligned} \|\mathbf{v} - \text{proj}_{\mathbf{d}}(\mathbf{v})\| &= \tfrac{1}{2} \left\| \begin{bmatrix} -3 \\ -3 \\ 2 \end{bmatrix} \right\| \\ &= \tfrac{1}{2}\sqrt{9 + 9 + 4} \\ &= \tfrac{1}{2}\sqrt{22} \end{aligned}$$

◆

Observação: ◆ Em termos de nossa notação anterior, d(B, r) = d($\mathbf{v}$, proj$_{\mathbf{d}}$($\mathbf{v}$)).

Quando a reta r está em $\mathbb{R}^2$ e sua equação geral é $ax + by = c$, a distância d(B, r) de $B = (x_0, y_0)$ à reta é dada pela fórmula

$$d(B, r) = \frac{|ax_0 + by_0 - c|}{\sqrt{a^2 + b^2}} \tag{3}$$

Você está convidado a demonstrar essa fórmula no Exercício 39.

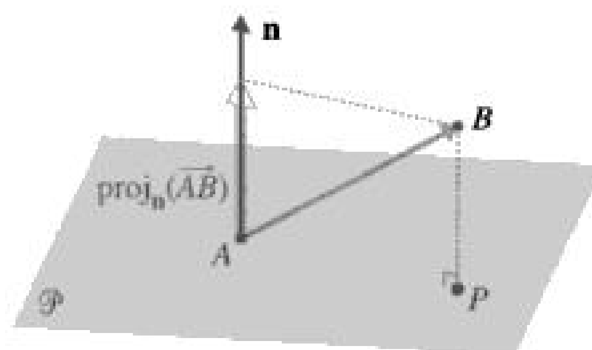


Figura 13 $d(B, \mathcal{P}) = \|\text{proj}_{\mathbf{n}}(\overrightarrow{AB})\|$

EXEMPLO 8 Determine a distância do ponto $B = (1, 0, 2)$ ao plano $\mathcal{P}$, cuja equação geral é $x + y - z = 1$.

SOLUÇÃO: Neste caso, precisamos calcular o comprimento de $\overrightarrow{PB}$, onde P é o ponto de $\mathcal{P}$ no pé da perpendicular a partir de B. Como mostra a Figura 13, se A for um ponto qualquer de $\mathcal{P}$ e aplicarmos o vetor $\mathbf{n} = \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix}$ normal a $\mathcal{P}$ de modo que sua origem seja em A, precisamos achar o comprimento da projeção de $\overrightarrow{AB}$ sobre $\mathbf{n}$.

Novamente, apresentamos os cálculos necessários em etapas.

Material com direitos autorais

Etapa 1. Por tentativa e erro, encontramos um ponto qualquer cujas coordenadas satisfaçam a equação $x + y - z = 1$. $A = (1, 0, 0)$ é um deles.

Etapa 2. Seja

$$\mathbf{v} = \overrightarrow{AB} = \mathbf{b} - \mathbf{a} = \begin{bmatrix} 1 \\ 0 \\ 2 \end{bmatrix} - \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 2 \end{bmatrix}$$

Etapa 3. A projeção de **v** sobre **n** é

$$\begin{aligned} \text{proj}_{\mathbf{n}}(\mathbf{v}) &= \left(\frac{\mathbf{n} \cdot \mathbf{v}}{\mathbf{n} \cdot \mathbf{n}}\right)\mathbf{n} \\ &= \left(\frac{1 \cdot 0 + 1 \cdot 0 - 1 \cdot 2}{1 + 1 + (-1)^2}\right)\begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix} \\ &= -\tfrac{2}{3}\begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix} \\ &= \begin{bmatrix} -\tfrac{2}{3} \\ -\tfrac{2}{3} \\ \tfrac{2}{3} \end{bmatrix} \end{aligned}$$

Etapa 4. A distância $d(B, \mathcal{P})$ de B a $\mathcal{P}$ é

$$\begin{aligned} \|\text{proj}_{\mathbf{n}}(\mathbf{v})\| &= \left|-\tfrac{2}{3}\right| \left\|\begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix}\right\| \\ &= \tfrac{2}{3}\left\|\begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix}\right\| \\ &= \tfrac{2}{3}\sqrt{3} \end{aligned}$$

Em geral, a distância $d(B, \mathcal{P})$ do ponto $B = (x_0, y_0, z_0)$ ao plano cuja equação geral é $ax + by + cz = d$ é dada pela fórmula

$$d(B, \mathcal{P}) = \frac{|ax_0 + by_0 + cz_0 - d|}{\sqrt{a^2 + b^2 + c^2}} \tag{4}$$

Você será convidado a deduzir essa fórmula no Exercício 40. ◆

Material com direitos autorais

◆ EXERCÍCIOS 1.4 ◆

Nos Exercícios 1 e 2, escreva uma equação (a) na forma normal e (b) na forma geral para a reta que passa por P e possui vetor normal **n**.

1. $P = (0, 0), \mathbf{n} = \begin{bmatrix} 3 \\ 2 \end{bmatrix}$ **2.** $P = (1, 2), \mathbf{n} = \begin{bmatrix} 5 \\ -3 \end{bmatrix}$

Nos Exercícios de 3 a 6, escreva equações (a) na forma normal e (b) na forma paramétrica para a reta que passa por P e tem vetor diretor **d**.

3. $P = (1, 0), \mathbf{d} = \begin{bmatrix} -1 \\ 3 \end{bmatrix}$ **4.** $P = (-4, 4), \mathbf{d} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$

5. $P = (0, 0, 0), \mathbf{d} = \begin{bmatrix} 1 \\ -1 \\ 4 \end{bmatrix}$ **6.** $P = (3, 0, -2), \mathbf{d} = \begin{bmatrix} 0 \\ 2 \\ 5 \end{bmatrix}$

Nos Exercícios 7 e 8, escreva equações (a) na forma normal e (b) na forma geral para o plano que passa por P e possui vetor normal **n**.

7. $P = (0, 1, 0), \mathbf{n} = \begin{bmatrix} 3 \\ 2 \\ 1 \end{bmatrix}$

8. $P = (-3, 5, 1), \mathbf{n} = \begin{bmatrix} 1 \\ -1 \\ 5 \end{bmatrix}$

Nos Exercícios 9 e 10, escreva equações (a) na forma geral e (b) na forma paramétrica para o plano que passa por P e possui vetores diretores **u** *e* **v**.

9. $P = (0, 0, 0), \mathbf{u} = \begin{bmatrix} 2 \\ 1 \\ 2 \end{bmatrix}, \mathbf{v} = \begin{bmatrix} -3 \\ 2 \\ 1 \end{bmatrix}$

10. $P = (6, -4, -3), \mathbf{u} = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}, \mathbf{v} = \begin{bmatrix} -1 \\ 1 \\ 1 \end{bmatrix}$

Nos Exercícios 11 e 12, dê uma equação vetorial para a reta que passa por P e Q.

11. $P = (1, -2), Q = (3, 0)$

12. $P = (0, 1, -1), Q = (-2, 1, 3)$

Nos Exercícios 13 e 14, dê uma equação vetorial para o plano que passa pelos pontos P, Q e R.

13. $P = (1, 1, 1), Q = (4, 0, 2), R = (0, 1, -1)$

14. $P = (1, 0, 0), Q = (0, 1, 0), R = (0, 0, 1)$

15. Ache equações paramétricas e uma equação vetorial para as retas de $\mathbb{R}^2$ que possuem as seguintes equações:

(a) $y = 3x - 1$ (b) $3x + 2y = 5$

16. Considere a equação vetorial $\mathbf{x} = \mathbf{p} + t(\mathbf{q} - \mathbf{p})$, onde **p** e **q** correspondem a pontos distintos P e Q em $\mathbb{R}^2$ ou $\mathbb{R}^3$.

(a) Mostre que essa equação descreve o segmento de reta $\overline{PQ}$ quando t varia de 0 a 1.

(b) Para que valores de t **x** é ponto médio de $\overline{PQ}$? Qual é esse **x**?

(c) Determine o ponto médio de $\overline{PQ}$ quando $P = (2, -3)$ e $Q = (0, 1)$.

(d) Determine o ponto médio de $\overline{PQ}$ quando $P = (1, 0, 1)$ e $Q = (4, 1, -2)$.

(e) Determine os dois pontos que dividem o segmento $\overline{PQ}$ da parte (c) em três partes iguais.

(f) Determine os dois pontos que dividem o segmento $\overline{PQ}$ da parte (d) em três partes iguais.

17. Indique uma "demonstração vetorial" do fato de que, em $\mathbb{R}^2$, duas retas com inclinações m_1 e m_2 são perpendiculares se, e somente se, $m_1 m_2 = -1$.

18. A reta r passa pelo ponto $P = (1, -1, 1)$ e tem vetor diretor $\mathbf{d} = \begin{bmatrix} 2 \\ 3 \\ -1 \end{bmatrix}$. Para cada um dos seguintes planos $\mathscr{P}$, determine se r e $\mathscr{P}$ são paralelos, perpendiculares ou nenhum desses dois:

(a) $2x + 3y - z = 1$ (b) $4x - y + 5z = 0$

(c) $x - y - z = 3$ (d) $4x + 6y - 2z = 0$

19. O plano $\mathscr{P}_1$ tem equação $4x - y + 5z = 2$. Para cada um dos planos $\mathscr{P}$ do Exercício 18, determine se $\mathscr{P}_1$ e $\mathscr{P}$ são paralelos, perpendiculares ou nenhum desses dois.

20. Determine uma equação vetorial para a reta de $\mathbb{R}^2$ que passa por $P = (2, -1)$ e é perpendicular à reta de equação geral $2x - 3y = 1$.

21. Determine uma equação vetorial para a reta de $\mathbb{R}^2$ que passa por $P = (2, -1)$ e é paralela à reta de equação geral $2x - 3y = 1$.

22. Determine uma equação vetorial para a reta de $\mathbb{R}^3$ que passa por $P = (-1, 0, 3)$ e é perpendicular ao plano de equação geral $x - 3y + 2z = 5$.

23. Determine uma equação vetorial para a reta de $\mathbb{R}^3$ que passa por $P = (-1, 0, 3)$ e é perpendicular à reta de equações paramétricas

$$\begin{aligned} x &= 1 - t \\ y &= 2 + 3t \\ z &= -2 - t \end{aligned}$$

Material com direitos autorais

24. Escreva uma equação na forma normal para o plano que passa por $P = (0, -2, 5)$ e é paralelo ao plano de equação geral $6x - y + 2z = 3$.

25. Os vértices de um cubo são os oito pontos (x, y, z) tais que cada uma das coordenadas x, y e z são iguais a 0 ou 1. (Veja a Figura 10 na Seção 1.3.)

(a) Encontre uma equação geral para cada um dos planos que determinam as seis faces do cubo.

(b) Encontre uma equação geral para o plano que contém a diagonal que liga a origem ao ponto (1, 1, 1) e é perpendicular ao plano xy.

(c) Encontre uma equação geral para o plano que contém as diagonais das faces do cubo do Exemplo 8 da Seção 1.3.

26. Encontre uma equação para o conjunto de todos os pontos eqüidistantes dos pontos $P = (1, 0, -2)$ e $Q = (5, 2, 4)$.

Nos Exercícios 27 e 28, determine a distância do ponto Q à reta r.

27. $Q = (2, 2)$, r de equação $\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} -1 \\ 2 \end{bmatrix} + t\begin{bmatrix} 1 \\ -1 \end{bmatrix}$

28. $Q = (0, 1, 0)$, r de equação $\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} + t\begin{bmatrix} -2 \\ 0 \\ 3 \end{bmatrix}$

Nos Exercícios 29 e 30, determine a distância do ponto Q ao plano $\mathcal{P}$.

29. $Q = (2, 2, 2)$, $\mathcal{P}$ de equação $x + y - z = 0$

30. $Q = (0, 0, 0)$, $\mathcal{P}$ de equação $x - 2y + 2z = 1$

A Figura 14 sugere uma maneira de usar vetores para localizar o ponto R da reta r que está mais próximo de Q.

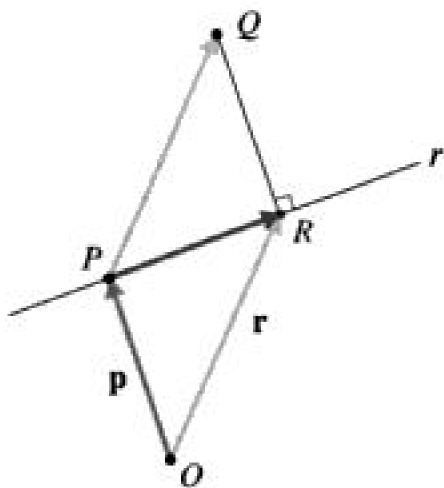


Figura 14 $\mathbf{r} = \mathbf{p} + \overrightarrow{PR}$

31. Encontre o ponto R de r mais próximo de Q no Exercício 27.

32. Encontre o ponto R de r mais próximo de Q no Exercício 28.

A Figura 15 sugere uma maneira de localizar o ponto R de $\mathcal{P}$ mais próximo de Q, usando vetores.

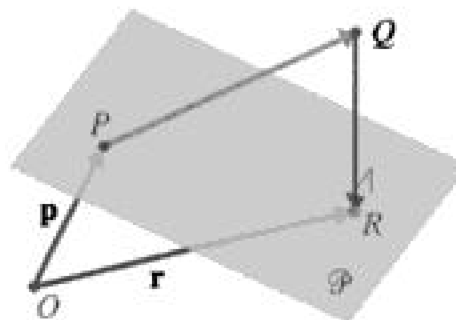


Figura 15 $\mathbf{r} = \mathbf{p} + \overrightarrow{PQ} + \overrightarrow{QR}$

33. Encontre o ponto R de $\mathcal{P}$ mais próximo de Q no Exercício 29.

34. Encontre o ponto R de $\mathcal{P}$ mais próximo de Q no Exercício 30.

Nos Exercícios 35 e 36, determine a distância entre as retas paralelas.

35. $\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix} + s\begin{bmatrix} -2 \\ 3 \end{bmatrix}$ e $\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 5 \\ 4 \end{bmatrix} + t\begin{bmatrix} -2 \\ 3 \end{bmatrix}$

36. $\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix} + s\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$ e $\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} + t\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$

Nos Exercícios 37 e 38, determine a distância entre os planos paralelos.

37. $2x + y - 2z = 0$ e $2x + y - 2z = 5$

38. $x + y + z = 1$ e $x + y + z = 3$

39. Prove a Equação 3 da página 41.

40. Prove a Equação 4 da página 42.

41. Prove que, em $\mathbb{R}^2$, a distância entre retas paralelas de equações $\mathbf{n} \cdot \mathbf{x} = c_1$ e $\mathbf{n} \cdot \mathbf{x} = c_2$ é dada por $\dfrac{|c_1 - c_2|}{\|\mathbf{n}\|}$.

42. Prove que a distância entre planos paralelos de equações $\mathbf{n} \cdot \mathbf{x} = d_1$ e $\mathbf{n} \cdot \mathbf{x} = d_2$ é dada por $\dfrac{|d_1 - d_2|}{\|\mathbf{n}\|}$.

Material com direitos autorais

Investigação

O Produto Vetorial

Seria conveniente se pudéssemos converter facilmente a equação vetorial do plano $\mathbf{x} = \mathbf{p} + s\mathbf{u} + t\mathbf{v}$ na forma normal $\mathbf{n} \cdot \mathbf{x} = \mathbf{n} \cdot \mathbf{p}$. Dados dois vetores não paralelos $\mathbf{u}$ e $\mathbf{v}$, o que precisamos é de um processo que produza um terceiro vetor $\mathbf{n}$ ortogonal a $\mathbf{u}$ e $\mathbf{v}$. Uma maneira é usar a operação conhecida como ***produto vetorial*** de vetores. *Válida somente em* $\mathbb{R}^3$, é assim definida:

Definição O *produto vetorial* de $\mathbf{u} = \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix}$ e $\mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix}$ é o vetor $\mathbf{u} \times \mathbf{v}$ definido por

$$\mathbf{u} \times \mathbf{v} = \begin{bmatrix} u_2v_3 - u_3v_2 \\ u_3v_1 - u_1v_3 \\ u_1v_2 - u_2v_1 \end{bmatrix}$$

Ilustramos a seguir um diagrama que pode ajudá-lo a lembrar como se calcula o produto vetorial de dois vetores. Sob cada vetor completo escrito em coluna, um ao lado do outro, escreva as duas primeiras coordenadas de cada vetor. Ignore as coordenadas da linha de cima e considere os blocos de quatro linhas. Subtraia os produtos das coordenadas ligadas por linhas pontilhadas dos produtos das coordenadas ligadas por linhas cheias. (Esse processo ajuda a perceber que a primeira coordenada de $\mathbf{u} \times \mathbf{v}$ não possui índices 1, a segunda não tem índices 2 e a terceira não tem índices 3.)

$$\begin{array}{ccc} u_1 & v_1 & \\ u_2 & v_2 & \\ u_3 & v_3 & u_2v_3 - u_3v_2 \\ u_1 & v_1 & u_3v_1 - u_1v_3 \\ u_2 & v_2 & u_1v_2 - u_2v_1 \end{array}$$

Os problemas a seguir exploram brevemente o produto vetorial.

1. Calcule $\mathbf{u} \times \mathbf{v}$.

(a) $\mathbf{u} = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}, \mathbf{v} = \begin{bmatrix} 3 \\ -1 \\ 2 \end{bmatrix}$ (b) $\mathbf{u} = \begin{bmatrix} 3 \\ -1 \\ 2 \end{bmatrix}, \mathbf{v} = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}$

(c) $\mathbf{u} = \begin{bmatrix} -1 \\ 2 \\ 3 \end{bmatrix}, \mathbf{v} = \begin{bmatrix} 2 \\ -4 \\ -6 \end{bmatrix}$ (d) $\mathbf{u} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \mathbf{v} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$

2. Mostre que $\mathbf{e}_1 \times \mathbf{e}_2 = \mathbf{e}_3$, $\mathbf{e}_2 \times \mathbf{e}_3 = \mathbf{e}_1$ e $\mathbf{e}_3 \times \mathbf{e}_1 = \mathbf{e}_2$.
3. Usando a definição de produto vetorial, prove que $\mathbf{u} \times \mathbf{v}$ é ortogonal a $\mathbf{u}$ e a $\mathbf{v}$.
4. Com a ajuda do produto vetorial, encontre uma equação na forma normal para o plano indicado a seguir:

Material com direitos autorais

(a) O plano que passa por $P = (1, 0, -2)$, é paralelo a $\mathbf{u} = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}$ e a $\mathbf{v} = \begin{bmatrix} 3 \\ -1 \\ 2 \end{bmatrix}$

(b) O plano que passa por $P = (0, -1, 1)$, $Q = (2, 0, 2)$ e $R = (1, 2, -1)$.

5. Prove as seguintes propriedades do produto vetorial:

(a) $\mathbf{v} \times \mathbf{u} = -(\mathbf{u} \times \mathbf{v})$
(b) $\mathbf{u} \times \mathbf{0} = \mathbf{0}$
(c) $\mathbf{u} \times \mathbf{u} = \mathbf{0}$
(d) $\mathbf{u} \times k\mathbf{v} = k(\mathbf{u} \times \mathbf{v})$
(e) $\mathbf{u} \times k\mathbf{u} = \mathbf{0}$
(f) $\mathbf{u} \times (\mathbf{v} + \mathbf{w}) = \mathbf{u} \times \mathbf{v} + \mathbf{u} \times \mathbf{w}$

6. Prove as seguintes propriedades do produto vetorial:

(a) $\mathbf{u} \cdot (\mathbf{v} \times \mathbf{w}) = (\mathbf{u} \times \mathbf{v}) \cdot \mathbf{w}$
(b) $\mathbf{u} \times (\mathbf{v} \times \mathbf{w}) = (\mathbf{u} \cdot \mathbf{w})\mathbf{v} - (\mathbf{u} \cdot \mathbf{v})\mathbf{w}$
(c) $\|\mathbf{u} \times \mathbf{v}\|^2 = \|\mathbf{u}\|^2\|\mathbf{v}\|^2 - (\mathbf{u} \cdot \mathbf{v})^2$

7. Refaça os problemas 2 e 3 usando os problemas 5 e 6.

8. Sejam $\mathbf{u}$ e $\mathbf{v}$ vetores em $\mathbb{R}^3$ e θ o ângulo entre $\mathbf{u}$ e $\mathbf{v}$.
(a) Prove que $\|\mathbf{u} \times \mathbf{v}\| = \|\mathbf{u}\| \, \|\mathbf{v}\| \operatorname{sen}\theta$. [*Sugestão*: consulte o Exercício 6(c).]
(b) Prove que a área $\mathcal{A}$ do triângulo determinado por $\mathbf{u}$ e $\mathbf{v}$ (como mostra a Figura 1) é dada por

$$\mathcal{A} = \tfrac{1}{2}\|\mathbf{u} \times \mathbf{v}\|$$

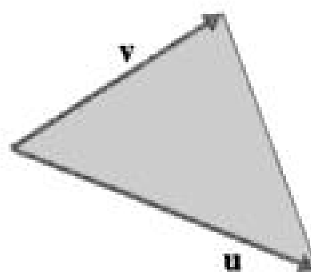


Figura 1

(c) Use o resultado da parte (b) para determinar a área do triângulo de vértices $A = (1, 2, 1)$, $B = (2, 1, 0)$ e $C = (5, -1, 3)$.

1.5 Vetores de Código e Aritmética Modular

A História nos mostra que as pessoas sempre transmitiram informações por meio de códigos. Por vezes, a intenção é encobrir a mensagem que está sendo enviada, como no caso em que cada letra de uma palavra é trocada por outra diferente, de acordo com uma regra de substituição. Embora fascinantes, esses códigos secretos ou cifras não são de nosso interesse aqui: são objetos de estudo do campo da *criptografia*. Em vez disso, falaremos sobre códigos utilizados quando a informação deve ser transmitida eletronicamente.

Um exemplo familiar de um desses códigos é o código Morse, com seu sistema de pontos e traços. O advento dos computadores digitais do século XX levou à necessidade de transmitir rapidamente e com precisão quantidades maciças de dados. Computadores foram projetados para codificar dados como seqüências de 0s e 1s. Muitos avanços tecnológicos recentes dependem de códigos, e nós os encontramos todos os dias sem nos darmos conta disso: comunicações via satélite, CD, Códigos Universais de Produtos (UPC – *Universal Product Codes*) associados aos códigos de barras encontrados em mercadorias e o Padrão Internacional de Numeração de Livros (ISBN – *International Standard Book Numbers*), encontrado em todo livro publicado são alguns exemplos.

Nesta seção, usaremos vetores para projetar códigos detectores de erros que podem ocorrer na transmissão de dados. Nos capítulos posteriores, construiremos códigos que não somente detectam, mas corrigem erros. Os vetores que aparecem no estudo de códigos não são os vetores familiares de $\mathbb{R}^n$, mas vetores com apenas

Material com direitos autorais

um número finito de possibilidades para cada componente. Tais vetores dependem de um tipo diferente de aritmética – chamada *aritmética modular* – que será apresentada nesta seção e utilizada por todo o livro.

Códigos Binários

A moderna teoria de códigos originou-se com o trabalho do matemático e cientista de computação americano Claude Shannon (1916–2001), cuja tese de 1937 mostrou como a álgebra podia desempenhar papel importante no projeto e análise de circuitos elétricos. Mais tarde, Shannon teria um papel decisivo na criação da teoria da informação e daria a base teórica para os hoje chamados códigos corretores de erro.

Já que os computadores descrevem dados em termos de 0s e 1s (que podem ser interpretados como desligado/ligado, fechado/aberto, falso/verdadeiro ou não/sim), começamos considerando códigos *binários*, que consistem em vetores cujas coordenadas são iguais a 0 ou 1. Nesse contexto, as regras usuais da aritmética devem ser modificadas, pois o resultado de cada cálculo que envolve escalares deve ser 0 ou 1. Veja a seguir as regras modificadas para adição e multiplicação.

+	0	1
0	0	1
1	1	0

·	0	1
0	0	0
1	0	1

A única curiosidade aqui é o fato de que $1 + 1 = 0$. Isto não é tão estranho quanto parece: se substituirmos 0 pela palavra "par" e 1 pela palavra "ímpar", essas tabelas simplesmente resumem as *regras de paridade* para a adição e multiplicação de inteiros pares e ímpares. Por exemplo, $1 + 1 = 0$ expressa a propriedade de que a soma de dois inteiros ímpares é um inteiro par. Com tais operações, nosso conjunto de escalares $\{0, 1\}$ é denotado por $\mathbb{Z}_2$ e é chamado de conjunto dos ***inteiros módulo 2***.

EXEMPLO 1 Em $\mathbb{Z}_2$, $1 + 1 + 0 + 1 = 1$ e $1 + 1 + 1 + 1 = 0$. (Esses cálculos ilustram novamente as regras de paridade: a soma de três ímpares e um par é ímpar; a soma de quatro ímpares é par.) ◆

Com $\mathbb{Z}_2$ como conjunto de escalares, podemos estender para os vetores as regras vistas há pouco. O conjunto de todas as n-uplas de 0s e 1s (com toda a aritmética módulo 2) é denotado por $\mathbb{Z}_2^n$. Os vetores em $\mathbb{Z}_2^n$ são chamados ***vetores binários de comprimento n***.[6]

EXEMPLO 2 Os vetores em $\mathbb{Z}_2^2$ são $[0, 0]$, $[0, 1]$, $[1, 0]$ e $[1, 1]$. (Em geral, quantos vetores possui $\mathbb{Z}_2^n$?) ◆

EXEMPLO 3 Sejam $\mathbf{u} = [1, 1, 0, 1, 0]$ e $\mathbf{v} = [0, 1, 1, 1, 0]$ dois vetores binários de comprimento 5. Determine $\mathbf{u} \cdot \mathbf{v}$.

SOLUÇÃO: O cálculo de $\mathbf{u} \cdot \mathbf{v}$ é feito em $\mathbb{Z}_2$, por isso temos

$$\begin{aligned}\mathbf{u} \cdot \mathbf{v} &= 1 \cdot 0 + 1 \cdot 1 + 0 \cdot 1 + 1 \cdot 1 + 0 \cdot 0 \\ &= 0 + 1 + 0 + 1 + 0 \\ &= 0\end{aligned}$$

◆

[6] Estamos usando o termo comprimento diferentemente daquele usado em $\mathbb{R}^n$. Isso não deve causar confusão, já que não existe nenhuma noção geométrica de comprimento para vetores binários.

Material com direitos autorais

Na prática, temos uma mensagem (em palavras, números ou símbolos) que queremos enviar. Começamos codificando cada "palavra" da mensagem por um vetor binário.

Definição Um ***código binário*** é um conjunto de vetores binários (de mesmo comprimento) chamados ***vetores de código***. O processo de conversão de uma mensagem em vetores de código é chamado ***codificação***, e o processo inverso é chamado ***decodificação***.

Como veremos, é altamente desejável que um código tenha outras propriedades, como a habilidade de detectar quando ocorre um erro na transmissão de um vetor de código, e, se possível, sugerir como corrigir o erro.

Códigos Detectores de Erro

Suponha que tenhamos codificado uma mensagem como um conjunto de vetores de código binários. Queremos enviar tais vetores através de um *canal* (como um transmissor de rádio, uma linha de telefone, um cabo de fibra ótica ou um CD a *laser*). Infelizmente, o canal pode ser "ruidoso" (por conta de interferência elétrica, sinais de interferência ou sujeira e arranhões). Como conseqüência, podem aparecer erros: alguns 0s podem ser trocados por 1s, e vice-versa. Como podemos nos precaver desse problema?

EXEMPLO 4 Desejamos codificar e transmitir uma mensagem que consiste em uma das palavras *up*, *down*, *left* ou *right*. Decidimos usar os quatro vetores de $\mathbb{Z}_2^2$ como nosso código binário, de acordo com a Tabela 1.

Se o receptor também tiver essa tabela e a mensagem codificada for transmitida sem erros, decodificar será trivial. No entanto, vamos admitir que tenha ocorrido um erro simples. (Por erro simples, entendemos uma alteração em exatamente uma coordenada do vetor de código.) Por exemplo, suponha o envio da mensagem "down", codificada como [0, 1], com um engano na transmissão da primeira coordenada em que o 0 foi trocado por 1. O receptor verá então a mensagem [1, 1] e a decodificará como "right". (Vamos apenas nos preocupar com os casos de um erro simples como este. Na prática, trabalha-se com a hipótese de que a probabilidade de ocorrer erros múltiplos é muito pequena.) Mesmo que o receptor percebesse (de alguma maneira) que ocorreu um erro simples, ele não saberia se o vetor de código correto era [0, 1] ou [1, 0].

Tabela 1

Mensagem	Up	Down	Left	Right
Código	[0, 0]	[0, 1]	[1, 0]	[1, 1]

Tabela 2

Mensagem	Up	Down	Left	Right
Código	[0, 0, 0]	[0, 1, 1]	[1, 0, 1]	[1, 1, 0]

Suponha o envio da mensagem com um código que seja subconjunto de $\mathbb{Z}_2^3$ – em outras palavras, um código binário de comprimento 3, como mostra a Tabela 2.

Esse código poderá detectar qualquer erro simples. Por exemplo, se "down" for enviado como [0, 1, 1] e ocorrer um erro em alguma coordenada, o receptor lerá [1, 1, 1], [0, 0, 1] ou [0, 1, 0], e nenhum deles é um vetor de código. Assim, ele saberá que ocorreu um erro (mas não onde) e poderá pedir que a mensagem codificada seja retransmitida. (Por que o receptor não saberá onde está o erro?) ◆

Material com direitos autorais

O código da Tabela 2 é um exemplo de ***código detector de erros***. Até os anos 40, isso era o melhor que podia ser obtido. O advento dos computadores digitais conduziu ao desenvolvimento de códigos que podiam *corrigir* tão bem quanto detectar erros. Consideraremos tais códigos nos Capítulos 3, 6 e 7.

A própria mensagem a ser transmitida pode ser formada por vetores binários. Neste caso, um código detector de erros simples mas muito usado é o ***código de checagem de paridade***, que é obtido anexando-se uma nova componente – chamada ***dígito de checagem*** – a cada vetor, de maneira que a paridade (o número total de 1s) seja um número par.

EXEMPLO 5 Se a mensagem a ser enviada for o vetor binário [1, 0, 0, 1, 0, 1], que possui um número ímpar de 1s, o dígito de checagem será 1 (de modo que o total de 1s do vetor de código seja par) e o vetor de código será [1, 0, 0, 1, 0, 1, 1].

Note que qualquer erro simples será detectado, pois ele mudará a paridade do vetor de código de par para ímpar. Por exemplo, se ocorrer um erro na terceira coordenada, o vetor de código será recebido como [1, 0, 1, 1, 0, 1, 1], cuja paridade é ímpar, pois tem cinco 1s. ◆

O termo *paridade* vem da palavra latina *par*. Dois inteiros possuem a mesma paridade se forem ambos pares ou ambos ímpares.

Vamos estudar esse conceito um pouco mais formalmente. Suponha que a mensagem seja o vetor binário $\mathbf{b} = [b_1, b_2, ..., b_n]$ em $\mathbb{Z}_2^n$. Então, o vetor código de checagem de paridade é $\mathbf{v} = [b_1, b_2, ..., b_n, d]$ em $\mathbb{Z}_2^{n+1}$, onde o dígito de checagem d é escolhido de maneira que

$$b_1 + b_2 \cdots + b_n + d = 0 \text{ em } \mathbb{Z}_2$$

ou, equivalentemente,

$$\mathbf{1} \cdot \mathbf{v} = 0$$

onde $\mathbf{1} = [1, 1, \ldots, 1]$, um vetor com todas as coordenadas iguais a 1. O vetor $\mathbf{1}$ é chamado ***vetor de checagem***. Se um vetor $\mathbf{v}'$ é recebido e $\mathbf{1} \cdot \mathbf{v}' = 1$, podemos estar certos de que ocorreu um erro. (Embora não estejamos considerando a possibilidade de mais do que um erro, observe que esse esquema não detectará uma quantidade par de erros.)

Os códigos de checagem de paridade são um caso especial de ***códigos de checagem de dígitos*** mais gerais, os quais serão considerados depois de estendermos as idéias precedentes a situações mais abrangentes.

Aritmética Modular

É possível estender o que acabamos de fazer para vetores binários a vetores cujas componentes são elementos do conjunto finito $\{0, 1, 2, \ldots, k\}$ para $k \geq 2$. Para tanto, devemos primeiro estender a idéia da aritmética binária.

EXEMPLO 6 Os ***inteiros módulo 3*** consistem no conjunto $\mathbb{Z}_3 = \{0, 1, 2\}$, com adição e multiplicação dadas assim:

+	0	1	2
0	0	1	2
1	1	2	0
2	2	0	1

·	0	1	2
0	0	0	0
1	0	1	2
2	0	2	1

Observe que o resultado de cada adição e multiplicação pertence ao conjunto {0, 1, 2}; dizemos que $\mathbb{Z}_3$ é ***fechado*** em relação às operações de adição e multiplicação. Talvez seja mais fácil pensar nesse conjunto em termos de um relógio de três horas, com 0, 1 e 2 no seu mostrador, como mostra a Figura 1.

Material com direitos autorais

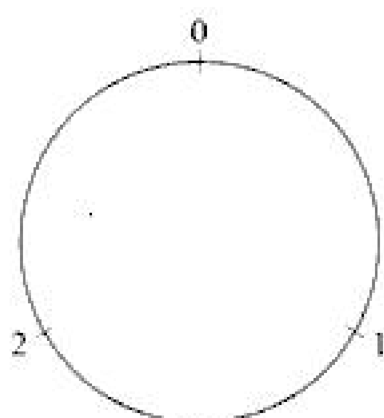


Figura 1 Aritmética módulo 3

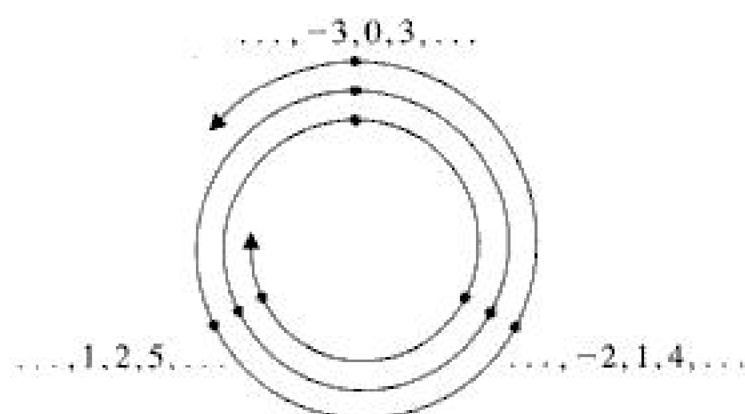


Figura 2

O cálculo $1 + 2 = 0$ é interpretado do seguinte modo: duas horas após 1 hora é 0 hora. Assim, como 24:00 e 12:00 são representados pela mesma marca em um relógio de doze horas, 3 e 0 são equivalentes em um relógio de três horas. Da mesma maneira, todos os múltiplos de 3 — positivos e negativos — são equivalentes a 0; o número 1 é equivalente a qualquer número que seja igual a 1 mais um múltiplo de 3 (como −2, 4 e 7); e 2 é equivalente a qualquer número que seja igual a 2 mais um múltiplo de 3 (como −1, 5 e 8). Podemos imaginar a reta numerada enrolada em volta de um círculo, como mostra a Figura 2. ◆

EXEMPLO 7 O número 3.548 é equivalente a que elemento de $\mathbb{Z}_3$?

SOLUÇÃO: Isso é o mesmo que perguntar onde está 3.548 no relógio de três horas. A idéia é calcular quão longe esse número está do múltiplo de 3 mais próximo e menor do que ele; isto é, precisamos *saber* o resto da divisão de 3.548 por 3. Efetuada a divisão, encontramos $3.548 = 3 \cdot 1182 + 2$, e, portanto, o resto é 2. Assim, 3.548 é equivalente a 2 em $\mathbb{Z}_3$. ◆

Em cursos de álgebra abstrata e teoria dos números, que exploram esse conceito com detalhes, a equivalência que acabamos de ver é freqüentemente escrita como $3.548 = 2 \pmod 3$ ou $3.548 \equiv 2 \pmod 3$, onde $\equiv$ se lê "é congruente a". Não usaremos essa notação ou terminologia aqui.

EXEMPLO 8 Em $\mathbb{Z}_3$, calcule $2 + 2 + 1 + 2$.

SOLUÇÃO 1: Usamos as mesmas idéias do Exemplo 7. A soma usual é $2 + 2 + 1 + 2 = 7 = 1 + 6$, e, portanto, sua divisão por 3 deixa resto 1. Assim, $2 + 2 + 1 + 2 = 1$ em $\mathbb{Z}_3$.

SOLUÇÃO 2: Um jeito melhor de efetuar esse cálculo é fazê-lo passo a passo em $\mathbb{Z}_3$.

$$\begin{aligned}2 + 2 + 1 + 2 &= (2 + 2) + 1 + 2 \\ &= 1 + 1 + 2 \\ &= (1 + 1) + 2 \\ &= 2 + 2 \\ &= 1\end{aligned}$$

Usamos parênteses para agrupar os termos que escolhemos para somar. Poderíamos acelerar os cálculos combinando simultaneamente os dois primeiros e os dois últimos termos:

$$\begin{aligned}(2 + 2) + (1 + 2) &= 1 + 0 \\ &= 1\end{aligned}$$

Multiplicações repetidas podem ser efetuadas analogamente. A idéia é usar as tábuas de adição e multiplicação para reduzir o resultado de cada conta a 0, 1 ou 2. ◆

Material com direitos autorais

A extensão dessas idéias para vetores é direta.

EXEMPLO 9 Em $\mathbb{Z}_3^5$, sejam $\mathbf{u} = [2, 2, 0, 1, 2]$ e $\mathbf{v} = [1, 2, 2, 2, 1]$. Então:

$$\begin{aligned}\mathbf{u}\cdot\mathbf{v} &= 2\cdot 1 + 2\cdot 2 + 0\cdot 2 + 1\cdot 2 + 2\cdot 1\\ &= 2 + 1 + 0 + 2 + 2\\ &= 1\end{aligned}$$

Vetores em $\mathbb{Z}_3^5$ são chamados ***vetores ternários de comprimento* 5**. ◆

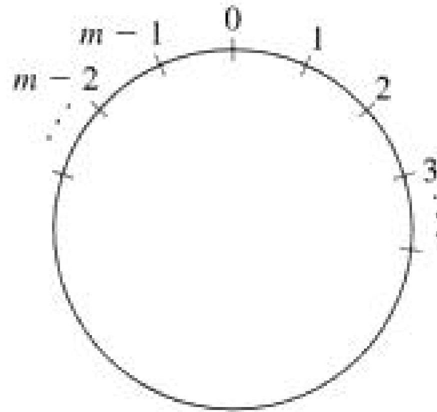


Figura 3 Aritmética módulo m

Em geral, temos o conjunto $\mathbb{Z}_m = \{0, 1, 2, \ldots, m-1\}$ de ***inteiros módulo m*** (correspondente a um relógio de m horas, como mostra a Figura 3) e o conjunto dos ***vetores m-ários de comprimento n***, denotado por $\mathbb{Z}_m^n$. Códigos que utilizam vetores m-ários são chamados ***códigos m-ários***.

O próximo exemplo é uma extensão direta do Exemplo 5 para códigos ternários.

EXEMPLO 10 Seja $\mathbf{b} = [b_1, b_2, \ldots, b_n]$ um vetor em $\mathbb{Z}_3^n$. Então, um vetor código de checagem de dígito pode ser definido por $\mathbf{v} = [b_1, b_2, \ldots, b_n, d]$ (em $\mathbb{Z}_3^{n+1}$), com o dígito de checagem d escolhido de maneira que

$$\mathbf{1}\cdot\mathbf{v} = 0$$

(onde o vetor de checagem $\mathbf{1} = [1, 1, \ldots, 1]$ é o vetor de 1s em $\mathbb{Z}_3^{n+1}$). Portanto, o dígito de checagem satisfaz

$$b_1 + b_2 + \ldots + b_n + d = 0 \text{ em } \mathbb{Z}_3$$

Por exemplo, considere o vetor $\mathbf{u} = [2, 2, 0, 1, 2]$ do Exemplo 9. A soma de suas coordenadas é $2 + 2 + 0 + 1 + 2 = 1$, e, portanto, o dígito de checagem deve ser 2 (pois $1 + 2 = 0$). Assim, o vetor código associado é $\mathbf{v} = [2, 2, 0, 1, 2, 2]$. ◆

Embora códigos simples de checagem de dígitos detectem erros simples, freqüentemente é importante detectar outros tipos comuns de erros, como uma troca acidental, ou *transposição*, de duas coordenadas adjacentes. (Por exemplo, a transposição da segunda e terceira coordenadas de $\mathbf{v}$ no Exemplo 10 irá resultar no vetor incorreto $\mathbf{v}' = [2, 0, 2, 1, 2, 2]$.) Para tais problemas, outros tipos de códigos de checagem de dígitos são projetados. Muitos deles simplesmente substituem o vetor de checagem $\mathbf{1}$ por algum outro vetor $\mathbf{c}$ cuidadosamente escolhido.



Figura 4 Um UPC

EXEMPLO 11 O UPC (Figura 4) é um código associado aos códigos de barra encontrados em muitos tipos de mercadorias.

Material com direitos autorais

As barras pretas e brancas escaneadas pelo laser da caixa registradora do mercado correspondem a um vetor 10-ário* $\mathbf{v} = [\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_{11}, d]$ de comprimento 12. As 11 primeiras componentes constituem um vetor em $\mathbb{Z}_{10}^{11}$, que dão informação sobre o fabricante e o produto; a última componente d é um dígito de checagem escolhido de maneira que $\mathbf{c} \cdot \mathbf{v} = 0$ em $\mathbb{Z}_{10}$, onde o vetor de checagem $\mathbf{c}$ é dado por $\mathbf{c} = [3, 1, 3, 1, 3, 1, 3, 1, 3, 1, 3, 1]$. Portanto, depois de um rearranjo, temos

$$3(\mathbf{v}_1 + \mathbf{v}_3 + \mathbf{v}_5 + \mathbf{v}_7 + \mathbf{v}_9 + \mathbf{v}_{11}) + (\mathbf{v}_2 + \mathbf{v}_4 + \mathbf{v}_6 + \mathbf{v}_8 + \mathbf{v}_{10}) + d = 0$$

onde d é o dígito de checagem. Em outras palavras, o dígito de checagem é escolhido de maneira que o lado esquerdo da expressão seja um múltiplo de 10.

Para o UPC mostrado na Figura 4, podemos determinar que o dígito de checagem é 6, fazendo todas as contas em $\mathbb{Z}_{10}$:

$$\begin{aligned}
\mathbf{c} \cdot \mathbf{v} &= 3 \cdot 0 + 7 + 3 \cdot 4 + 9 + 3 \cdot 2 + 7 + 3 \cdot 0 + 2 + 3 \cdot 0 + 9 + 3 \cdot 4 + d \\
&= 3(0 + 4 + 2 + 0 + 0 + 4) + (7 + 9 + 7 + 2 + 9) + d \\
&= 3(0) + 4 + d \\
&= 4 + d
\end{aligned}$$

O dígito de checagem d deve ser 6 para se obter o total 0 dos cálculos em $\mathbb{Z}_{10}$.

(Uma outra maneira de pensar no dígito de checagem, neste exemplo, é que ele é escolhido de maneira que $\mathbf{c} \cdot \mathbf{v}$ seja um *múltiplo* de 10.) ◆

O UPC detectará todos os erros simples e a maior parte dos erros de transposição de coordenadas adjacentes. Para exemplificar este último ponto, suponha que o UPC no Exemplo 11 seja incorretamente escrito como $\mathbf{v}' = [0, 7, 4, 2, 9, 7, 0, 2, 0, 9, 4, 6]$, com a quarta e a quinta coordenadas em posições trocadas. Quando aplicarmos o vetor de checagem, teremos $\mathbf{c} \cdot \mathbf{v}' = 4 \neq 0$ alertando-nos para o fato de que deve ter havido um erro. (Veja os Exercícios 48 e 49.)

EXEMPLO 12 O ISBN é um outro código de checagem de dígito amplamente usado. É projetado para detectar mais tipos de erros que o Código Universal de Produto e, conseqüentemente, é um pouco mais complicado. O princípio básico ainda é o mesmo.

O vetor de código é um vetor em $\mathbb{Z}_{11}^{10}$. As primeiras nove componentes dão país, editor e informações sobre o livro; a décima componente é o dígito de checagem. O ISBN para o livro *Calculus: Concepts and contexts*, de James Stewart, é 0–534–34450–X. É descrito como o vetor

$$\mathbf{b} = [0, 5, 3, 4, 3, 4, 4, 5, 0, X]$$

onde o "dígito" de checagem é a letra X.

Para o código ISBN, o vetor de checagem é o vetor $\mathbf{c} = [10, 9, 8, 7, 6, 5, 4, 3, 2, 1]$, e a condição é que $\mathbf{c} \cdot \mathbf{b} = 0$ em $\mathbb{Z}_{11}$. Vamos determinar o dígito de checagem para o vetor $\mathbf{b}$ neste exemplo. Devemos calcular

$$\mathbf{c} \cdot \mathbf{b} = 10 \cdot 0 + 9 \cdot 5 + 8 \cdot 3 + 7 \cdot 4 + 6 \cdot 3 + 5 \cdot 4 + 4 \cdot 4 + 3 \cdot 5 + 2 \cdot 0 + d$$

onde d é o dígito de checagem. Começamos efetuando todas as multiplicações em $\mathbb{Z}_{11}$. (Por exemplo, $9 \cdot 5 = 1$, pois 45 é 1 mais o múltiplo mais próximo e menor que 45 – isto é, 44. Em um relógio de onze horas, 45 horas é 1 hora.) A soma simplificada é

$$0 + 1 + 2 + 6 + 7 + 9 + 5 + 4 + 0 + d$$

e, somando-se em $\mathbb{Z}_{11}$, obtemos $1 + d$. Agora, o dígito de checagem d deve ser escolhido de maneira que o

* N.T.: Também chamado decimal, ou seja, cujas coordenadas pertencem ao conjunto $\{0, 1, \ldots, 9\}$

Material com direitos autorais

resultado final seja 0; portanto, em $\mathbb{Z}_{11}$, $d = 10$. (Equivalentemente, d deve ser escolhido de modo que $\mathbf{c} \cdot \mathbf{b}$ seja múltiplo de 11.) Como é preferível que cada componente de um ISBN se componha de um único dígito, é usado o numeral romano X no lugar de 10 toda vez que ele ocorre como um dígito de checagem. ◆

O código ISBN detectará todos os erros simples e erros de transposição (veja os Exercícios 52-54).

◆ EXERCÍCIOS 1.5 ◆

Nos Exercícios de 1 a 4, **u** *e* **v** *são vetores binários. Determine* **u** + **v** *e* **u** · **v** *em cada caso.*

1. $\mathbf{u} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \mathbf{v} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ **2.** $\mathbf{u} = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \mathbf{v} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$

3. $\mathbf{u} = [1, 0, 1, 1], \mathbf{v} = [1, 1, 1, 1]$

4. $\mathbf{u} = [1, 1, 0, 1, 0], \mathbf{v} = [0, 1, 1, 1, 0]$

5. Escreva as tabelas de adição e multiplicação em $\mathbb{Z}_4$.

6. Escreva as tabelas de adição e multiplicação em $\mathbb{Z}_5$.

Nos Exercícios de 7 a 19, efetue os cálculos indicados.

7. $2 + 2 + 2$ em $\mathbb{Z}_3$ **8.** $2 \cdot 2 \cdot 2$ em $\mathbb{Z}_3$

9. $2(2 + 1 + 2)$ em $\mathbb{Z}_3$ **10.** $3 + 1 + 2 + 3$ em $\mathbb{Z}_4$

11. $2 \cdot 3 \cdot 2$ em $\mathbb{Z}_4$ **12.** $3(3 + 3 + 2)$ em $\mathbb{Z}_4$

13. $2 + 1 + 2 + 2 + 1$ em $\mathbb{Z}_3$, $\mathbb{Z}_4$ e $\mathbb{Z}_5$

14. $(3 + 4)(3 + 2 + 4 + 2)$ em $\mathbb{Z}_5$

15. $8(6 + 4 + 3)$ em $\mathbb{Z}_9$ **16.** 2^{100} em $\mathbb{Z}_{11}$

17. $[2, 1, 2] + [2, 0, 1]$ em $\mathbb{Z}_3^3$

18. $[2, 1, 2] \cdot [2, 2, 1]$ em $\mathbb{Z}_3^3$

19. $[2, 0, 3, 2] \cdot ([3, 1, 1, 2] + [3, 3, 2, 1])$ em $\mathbb{Z}_4^4$ e em $\mathbb{Z}_5^4$

Nos Exercícios de 20 a 31, resolva a equação dada ou indique que não há solução.

20. $x + 3 = 2$ em $\mathbb{Z}_5$ **21.** $x + 5 = 1$ em $\mathbb{Z}_6$

22. $2x = 1$ em $\mathbb{Z}_3$ **23.** $2x = 1$ em $\mathbb{Z}_4$

24. $2x = 1$ em $\mathbb{Z}_5$ **25.** $3x = 4$ em $\mathbb{Z}_5$

26. $3x = 4$ em $\mathbb{Z}_6$ **27.** $6x = 5$ em $\mathbb{Z}_8$

28. $8x = 9$ em $\mathbb{Z}_{11}$ **29.** $2x + 3 = 2$ em $\mathbb{Z}_5$

30. $4x + 5 = 2$ em $\mathbb{Z}_6$ **31.** $6x + 3 = 1$ em $\mathbb{Z}_8$

32. (a) Para quais valores de a a equação $x + a = 0$ possui solução em $\mathbb{Z}_5$?

(b) Para quais valores de a e b a equação $x + a = b$ possui solução em $\mathbb{Z}_6$?

(c) Para quais valores de a, b e m a equação $x + a = b$ possui solução em $\mathbb{Z}_m$?

33. (a) Para quais valores de a a equação $ax = 1$ tem solução em $\mathbb{Z}_5$?

(b) Para quais valores de a a equação $ax = 1$ tem solução em $\mathbb{Z}_6$?

(c) Para que valores de a e m a equação $ax = 1$ tem solução em $\mathbb{Z}_m$?

Nos Exercícios 34 e 35, determine o vetor **v** *de código de checagem de paridade para cada vetor binário* **u**.

34. $\mathbf{u} = [1, 0, 1, 1]$ **35.** $\mathbf{u} = [1, 1, 0, 1, 1]$

Nos Exercícios de 36 a 39, é dado um vetor **v** *de código de checagem de paridade. Determine se pode ter ocorrido um erro na transmissão de* **v**.

36. $\mathbf{v} = [1, 0, 1, 0]$ **37.** $\mathbf{v} = [1, 1, 1, 0, 1, 1]$

38. $\mathbf{v} = [0, 1, 0, 1, 1, 1]$ **39.** $\mathbf{v} = [1, 1, 0, 1, 0, 1, 1, 1]$

Os Exercícios de 40 a 43 se referem a códigos de checagem de dígitos nos quais o vetor de checagem **c** *é o vetor* $\mathbf{1} = [1, 1, \ldots, 1]$ *de comprimento apropriado. Em cada caso, encontre o dígito de checagem que deve ser anexado ao vetor* **u**.

40. $\mathbf{u} = [1, 2, 2, 2]$ em $\mathbb{Z}_3^4$ **41.** $\mathbf{u} = [3, 4, 2, 3]$ em $\mathbb{Z}_5^4$

42. $\mathbf{u} = [1, 5, 6, 4, 5]$ em $\mathbb{Z}_7^5$ **43.** $\mathbf{u} = [3, 0, 7, 5, 6, 8]$ em $\mathbb{Z}_9^6$

44. Prove que, para quaisquer inteiros positivos m e n, o código de checagem de dígito em $\mathbb{Z}_m^n$, com vetor de checagem $\mathbf{c} = \mathbf{1} = [1, 1, \ldots, 1]$, detectará todos os erros simples. (Isto é, prove que, se vetores **u** e **v** em $\mathbb{Z}_m^n$ diferem em exatamente uma coordenada, então $\mathbf{c} \cdot \mathbf{u} \neq \mathbf{c} \cdot \mathbf{v}$.)

Nos Exercícios 45 e 46, encontre o dígito de checagem d do UPC.

45. $[0, 5, 9, 4, 6, 4, 7, 0, 0, 2, 7, d]$

46. $[0, 1, 4, 0, 1, 4, 1, 8, 4, 1, 2, d]$

47. Considere o UPC $[0, 4, 6, 9, 5, 6, 1, 8, 2, 0, 1, 5]$.

(a) Mostre que esse UPC não pode estar correto.

(b) Assumindo que um único erro foi cometido e que o dígito incorreto é o 6 na terceira coordenada, encontre o UPC correto.

Material com direitos autorais

48. Prove que o UPC detectará todos os erros simples.

49. (a) Prove que, se ocorrer um erro de transposição na segunda e terceira coordenadas do UPC [0, 7, 4, 9, 2, 7, 0, 2, 0, 9, 4, 6], tal erro será detectado.

(b) Mostre que existe uma transposição envolvendo duas coordenadas adjacentes do UPC na parte (a) que não seriam detectadas.

(c) Em geral, quando o UPC não detecta um erro de transposição que envolve duas coordenadas adjacentes?

Nos Exercícios 50 e 51, *encontre o dígito de checagem d do* ISBN *dado.*

50. [0, 3, 8, 7, 9, 7, 9, 9, 3, d]

51. [0, 3, 9, 4, 7, 5, 6, 8, 2, d]

52. Considere o ISBN [0, 4, 4, 9, 5, 0, 8, 3, 5, 6].

(a) Mostre que esse ISBN não pode estar correto.

(b) Assumindo que um erro simples foi feito e que o dígito incorreto é o 5 na quinta coordenada, encontre o ISBN correto.

53. (a) Prove que, se um erro de transposição for feito na quarta e quinta coordenadas do ISBN [0, 6, 7, 9, 7, 6, 2, 9, 0, 6], esse erro será detectado.

(b) Prove que, se um erro de transposição ocorrer em duas coordenadas adjacentes quaisquer do ISBN da parte (a), ele será detectado.

(c) Prove que, em geral, o código ISBN sempre detecta um erro de transposição que envolve duas coordenadas adjacentes.

54. Considere o ISBN [0, 8, 3, 7, 0, 9, 9, 0, 2, 6].

(a) Mostre que esse ISBN não pode estar correto.

(b) Assumindo que ocorreu um erro de transposição envolvendo duas coordenadas adjacentes, encontre o ISBN correto.

(c) Dê um exemplo de um ISBN para o qual um erro de transposição envolvendo duas coordenadas adjacentes será detectado, mas não corrigível.

Material com direitos autorais

2

Sistemas de Equações Lineares

O mundo estava cheio de equações...
Deve haver uma resposta para todas as coisas, se
pelo menos você souber como colocar as questões.

– Anne Tyler
The accidental tourist
Alfred A. Knopf, 1985, p. 235

2.1 Introdução: Trivialidade

A palavra *trivial* deriva das palavras latinas *tri* ("três") e *via* ("caminho"). Assim, literalmente, uma trivialidade se refere a um lugar onde três caminhos se encontram. Esse ponto de encontro comum dá origem a um outro significado, mais comum, para *trivial*: lugar-comum, ordinário, insignificante.

Nesta seção, começamos a examinar sistemas de equações lineares. O mesmo sistema de equações pode ser visto de três modos diferentes, mas igualmente importantes – esses serão nossos três caminhos, todos levando à mesma solução. Você precisará acostumar-se com essas três maneiras de interpretar sistemas de equações lineares para que isso se torne ordinário (trivial!) para você.

O sistema de equações que iremos considerar é

$$\begin{aligned} 2x + y &= 8 \\ x - 3y &= -3 \end{aligned}$$

Problema 1: Desenhe duas retas representadas por essas equações. Qual é o seu ponto de interseção?

Problema 2: Considere os vetores $\mathbf{u} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$ e $\mathbf{v} = \begin{bmatrix} 1 \\ -3 \end{bmatrix}$. Desenhe o sistema de coordenadas determinado por **u** e **v**. (*Sugestão*: desenhe, de leve, primeiro o sistema de coordenadas padrão e use-o como ajuda para o desenho do novo sistema.)

Problema 3: No sistema u–v, encontre as coordenadas de $\mathbf{w} = \begin{bmatrix} 8 \\ -3 \end{bmatrix}$.

Problema 4: Uma outra maneira de colocar o Problema 3 é pedir que se determinem as coordenadas x e y para as quais $x\mathbf{u} + y\mathbf{v} = \mathbf{w}$. Escreva as duas equações às quais essa equação vetorial é equivalente (uma equação para cada componente). O que você observa?

Problema 5: Retorne agora às retas que você desenhou no Problema 1. Vamos nos referir à reta de equação $2x + y = 8$ como reta 1, e à reta de equação $x - 3y = -3$ como reta 2. Marque o ponto $(0, 0)$ em seu gráfico do

Material com direitos autorais

Problema 1 e chame-o de P_0. Desenhe um segmento de reta *horizontal* de P_0 até encontrar a reta 1 e chame esse novo ponto de P_1. Em seguida, desenhe um segmento de reta *vertical* de P_1 até encontrar a reta 2 e chame esse novo ponto de P_2. Agora, desenhe um segmento de reta *horizontal* de P_2 até a reta 1, obtendo o ponto P_3. Continue desenhando segmentos verticais até encontrar a reta 2, seguidos de segmentos horizontais para a reta 1. O que parece estar acontecendo?

Problema 6: Usando uma calculadora com duas casas decimais de precisão, encontre as coordenadas (aproximadas) dos pontos $P_1, P_2, P_3, \ldots, P_6$. (Você achará útil primeiramente resolver a primeira equação para x em termos de y e a segunda equação para y em termos de x.) Registre seus resultados na Tabela 1, escrevendo as coordenadas x e y de cada ponto separadamente.

Tabela 1

Ponto	x	y
P_0	0	0
P_1		
P_2		
P_3		
P_4		
P_5		
P_6		

Os resultados desses problemas mostram que a tarefa de "resolver" um sistema de equações lineares pode ser vista de várias maneiras. Repita o processo descrito nos problemas com os seguintes sistemas de equações:

(a) $\begin{aligned} 4x - 2y &= 0 \\ x + 2y &= 5 \end{aligned}$ (b) $\begin{aligned} 3x + 2y &= 9 \\ x + 3y &= 10 \end{aligned}$ (c) $\begin{aligned} x + y &= 5 \\ x - y &= 3 \end{aligned}$ (d) $\begin{aligned} x + 2y &= 4 \\ 2x - y &= 3 \end{aligned}$

Todas as suas observações feitas nos problemas 1 a 6 são ainda válidas para esses exemplos? Repare em qualquer similaridade ou diferença. Neste capítulo, iremos explorar essas idéias com mais detalhe.

2.2 Introdução aos Sistemas de Equações Lineares

Lembre-se de que a equação geral de uma reta em $\mathbb{R}^2$ é da forma

$$ax + by = c$$

e que a equação geral de um plano em $\mathbb{R}^3$ é da forma

$$ax + by + cz = d$$

Equações como as acima são chamadas ***equações lineares***.

Material com direitos autorais

Definição Uma ***equação linear*** nas n variáveis $x_1, x_2, \ldots, x_n$ é uma equação que pode ser escrita na forma

$$a_1x_1 + a_2x_2 + \cdots + a_nx_n = b$$

onde os ***coeficientes*** $a_1, a_2, \ldots, a_n$ e o ***termo independente*** b são constantes.

EXEMPLO 1 As seguintes equações são lineares:

$$3x - 4y = -1 \qquad r - \tfrac{1}{2}s - \tfrac{15}{3}t = 9 \qquad x_1 + 5x_2 = 3 - x_3 + 2x_4$$
$$\sqrt{2}x + \frac{\pi}{4}y - \left(\text{sen}\frac{\pi}{5}\right)z = 1 \qquad 3{,}2x_1 - 0{,}01x_2 = 4{,}6$$

Observe que a terceira equação é linear porque pode ser reescrita na forma $x_1 + 5x_2 + x_3 - 2x_4 = 3$. É também importante notar que, embora nesses exemplos (e na maioria das aplicações) os coeficientes e termos constantes sejam números reais, em alguns exemplos e aplicações eles serão números complexos ou elementos de $\mathbb{Z}_p$ para algum número primo p.

As seguintes equações não são lineares:

$$xy + 2z = 1 \qquad x_1^2 - x_2^3 = 3 \qquad \frac{x}{y} + z = 2$$
$$\sqrt{2x} + \frac{\pi}{4}y - \text{sen}\left(\frac{\pi}{5}z\right) = 1 \qquad \text{sen}\,x_1 - 3x_2 + 2^{x_3} = 0$$

Equações lineares, portanto, não contêm produtos, recíprocos ou outras funções das variáveis; as variáveis aparecem somente na potência 1 e multiplicadas apenas por constantes. Preste atenção especial no quarto exemplo de cada lista: por que a quarta equação na primeira lista é linear, mas a quarta equação na segunda lista não é? ◆

Uma ***solução*** de uma equação linear $a_1x_1 + a_2x_2 + \ldots + a_n x_n = b$ é um vetor $[s_1, s_2, \ldots, s_n]$, cujas coordenadas satisfazem à equação quando substituímos $x_1 = s_1, x_2 = s_2, \ldots, x_n = s_n$.

EXEMPLO 2 (a) $[5, 4]$ é uma solução de $3x - 4y = -1$ porque, quando substituímos $x = 5$ e $y = 4$, a equação é satisfeita: $3(5) - 4(4) = -1$. O vetor $[1, 1]$ é uma outra solução. As soluções simplesmente correspondem aos pontos da reta determinada pela equação dada. Assim, colocando $x = t$ e resolvendo para y, vemos que o conjunto completo das soluções pode ser escrito na forma paramétrica $[t, \frac{1}{4} + \frac{3}{4}t]$. (Poderíamos também colocar y igual a algum parâmetro — digamos, s — e resolver a equação para x; as duas soluções paramétricas pareceriam diferentes, mas seriam equivalentes. Tente isso.)

(b) A equação linear $x_1 - x_2 + 2x_3 = 3$ tem $[3, 0, 0]$, $[0, 1, 2]$ e $[6, 1, -1]$ como soluções particulares. O conjunto completo das soluções corresponde ao conjunto dos pontos pertencentes ao plano determinado pela equação dada. Se colocarmos $x_2 = s$ e $x_3 = t$, uma solução paramétrica será dada por $[3 + s - 2t, s, t]$. (Quais valores de s e de t produzem as três soluções específicas apontadas?) ◆

Um ***sistema de equações lineares*** é um conjunto finito de equações lineares, cada uma com as mesmas variáveis. Uma ***solução*** de um sistema de equações lineares é um vetor que é *simultaneamente* solução de cada uma das equações do sistema. O ***conjunto solução*** de um sistema de equações lineares é o conjunto de *todas* as soluções do sistema. Vamos nos referir ao processo de encontrar o conjunto solução de um sistema de equações lineares como "resolver o sistema".

Material com direitos autorais

EXEMPLO 3 O sistema

$$\begin{aligned} 2x - y &= 3 \\ x + 3y &= 5 \end{aligned}$$

tem $[2, 1]$ como solução, pois esse vetor é uma solução de ambas as equações. Por outro lado, $[1, -1]$ não é uma solução do sistema, já que ele satisfaz apenas a primeira equação. ◆

EXEMPLO 4 Resolva os seguintes sistemas de equações lineares:

$$\text{(a)}\ \begin{aligned} x - y &= 1 \\ x + y &= 3 \end{aligned} \qquad \text{(b)}\ \begin{aligned} x - y &= 2 \\ 2x - 2y &= 4 \end{aligned} \qquad \text{(c)}\ \begin{aligned} x - y &= 1 \\ x - y &= 3 \end{aligned}$$

SOLUÇÃO:

(a) A soma das duas equações dá $2x = 4$, logo, $x = 2$, de onde concluímos que $y = 1$. Uma rápida conta confirma que $[2, 1]$ é na verdade uma solução de ambas as equações. O fato de essa ser a *única* solução pode ser confirmado quando observamos que ela corresponde ao (único) ponto de interseção $(2, 1)$ das retas com equações $x - y = 1$ e $x + y = 3$, como é mostrado na Figura 1(a). Portanto, $[2, 1]$ é uma *solução única.*

(b) A segunda equação nesse sistema é exatamente o dobro da primeira, por isso as soluções são as soluções da primeira equação apenas — a saber, os pontos da reta $x - y = 2$. Estes podem ser representados na forma paramétrica por $[2 + t, t]$. Assim, esse sistema tem *infinitas soluções* [Figura 1(b)].

(c) Dois números x e y não podem ter uma diferença de 1 e de 3 simultaneamente. Portanto, esse sistema *não tem soluções.* (Uma forma mais algébrica de aproximação pode ser subtrair a segunda equação da primeira, levando à absurda conclusão de que $0 = -2$.) Como a Figura 1(c) mostra, neste caso as retas correspondentes às equações são paralelas.

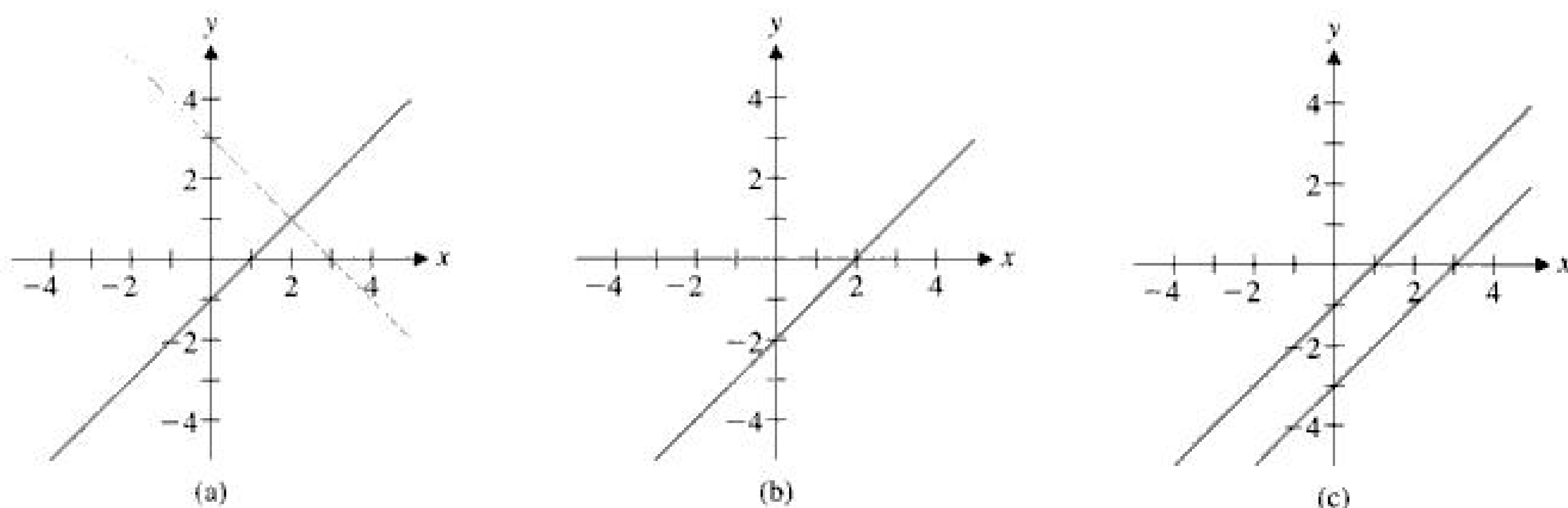


Figura 1 ◆

Um sistema de equações lineares é chamado de ***possível***** quando tem pelo menos uma solução. Um sistema sem nenhuma solução é chamado ***impossível***. Embora sejam pequenos, os três sistemas no Exemplo 4 ilustram as únicas três possibilidades para o número de soluções de um sistema de equações lineares com coeficientes reais. Provaremos mais tarde que essas mesmas três possibilidades valem para *qualquer* sistema de equações lineares sobre os números reais.

* N.T.: Em português, os termos *consistente* e *compatível* também são usados para nos referirmos a um sistema *possível*. Um sistema *impossível* também pode ser chamado de *inconsistente* ou *incompatível*.

Material com direitos autorais

Um sistema de equações lineares com coeficientes reais tem
(a) uma única solução (um sistema possível),
(b) infinitas soluções (um sistema possível) ou
(c) nenhuma solução (um sistema impossível).



Resolução de um Sistema de Equações Lineares

Dois sistemas lineares são chamados ***equivalentes*** quando têm os mesmos conjuntos solução. Por exemplo,

$$\begin{aligned} x - y &= 1 \\ x + y &= 3 \end{aligned} \quad \text{e} \quad \begin{aligned} x - y &= 1 \\ y &= 1 \end{aligned}$$

são equivalentes, já que ambos têm $[2, 1]$ como única solução. (Verifique.)

Nossa estratégia para a resolução de um sistema de equações lineares é transformar o sistema dado em um sistema equivalente que seja mais fácil de resolver. O formato triangular do segundo exemplo que acabamos de ver (no qual a segunda equação tem uma variável a menos que a primeira) é o que teremos como meta.

EXEMPLO 5 Resolva o sistema

$$\begin{aligned} x - y - z &= 2 \\ y + 3z &= 5 \\ 5z &= 10 \end{aligned}$$

SOLUÇÃO: Começando com a última equação e trabalhando de trás para a frente, encontramos sucessivamente $z = 2$, $y = 5 - 3(2) = -1$ e $x = 2 + (-1) + 2 = 3$. Portanto, a única solução é $[3, -1, 2]$. ◆

Esse procedimento usado para resolver o sistema do Exemplo 5 é chamado ***substituição de trás para a frente***.

Voltemos agora à estratégia geral de transformar um sistema dado em outro equivalente que possa ser resolvido facilmente por substituição de trás para a frente. Esse processo será descrito mais detalhadamente na próxima seção; por ora, vamos simplesmente observá-lo em ação em um único exemplo.

EXEMPLO 6 Resolva o sistema

$$\begin{aligned} x - y - z &= 2 \\ 3x - 3y + 2z &= 16 \\ 2x - y + z &= 9 \end{aligned}$$

SOLUÇÃO: Para transformar esse sistema em um que exiba a estrutura triangular do Exemplo 5, precisamos primeiramente eliminar a variável x das Equações 2 e 3. Observe que a subtração de múltiplos apropriados da Equação 1 das Equações 2 e 3 produz o resultado desejado. Em seguida, observe que estamos operando com os coeficientes, não com as variáveis, por isso podemos economizar alguma escrita se registrarmos apenas os coeficientes e os termos constantes na *matriz*

$$\left[\begin{array}{rrr|r} 1 & -1 & -1 & 2 \\ 3 & -3 & 2 & 16 \\ 2 & -1 & 1 & 9 \end{array}\right]$$

Material com direitos autorais

A palavra *matriz* deriva da palavra latina *mater*, que significa "mãe". Quando o sufixo *-iz* é acrescentado, o significado torna-se "útero". Assim como um útero envolve um feto, os colchetes de uma matriz envolvem seus elementos, e assim como o útero dá origem a um bebê, uma matriz gera certos tipos de funções, chamadas *transformações lineares*. Uma matriz com *m* linhas e *n* colunas é chamada matriz *m x n* (pronuncia-se "*m* por *n*").

na qual as três primeiras colunas contêm os coeficientes das variáveis na ordem, a última coluna contém os termos constantes e a barra vertical serve para nos lembrar do sinal de igual nas equações. Essa matriz é chamada de ***matriz completa*** do sistema.

Há várias maneiras de converter um sistema dado em um com a estrutura triangular que estamos buscando. Os passos que iremos usar aqui são os mais próximos ao método geral descrito na próxima seção. Realizaremos a seqüência de operações no sistema dado e simultaneamente na matriz completa. Começaremos eliminando o x das Equações 2 e 3.

$$\begin{aligned} x - y - z &= 2 \\ 3x - 3y + 2z &= 16 \\ 2x - y + z &= 9 \end{aligned} \qquad \left[\begin{array}{rrr|r} 1 & -1 & -1 & 2 \\ 3 & -3 & 2 & 16 \\ 2 & -1 & 1 & 9 \end{array}\right]$$

Subtraindo três vezes a primeira equação da segunda equação:

Subtraindo três vezes a primeira linha da segunda linha:

$$\begin{aligned} x - y - z &= 2 \\ 5z &= 10 \\ 2x - y + z &= 9 \end{aligned} \qquad \left[\begin{array}{rrr|r} 1 & -1 & -1 & 2 \\ 0 & 0 & 5 & 10 \\ 2 & -1 & 1 & 9 \end{array}\right]$$

Subtraindo duas vezes a primeira equação da terceira equação:

Subtraindo duas vezes a primeira linha da terceira linha:

$$\begin{aligned} x - y - z &= 2 \\ 5z &= 10 \\ y + 3z &= 5 \end{aligned} \qquad \left[\begin{array}{rrr|r} 1 & -1 & -1 & 2 \\ 0 & 0 & 5 & 10 \\ 0 & 1 & 3 & 5 \end{array}\right]$$

Trocando as Equações 2 e 3:

Trocando as linhas 2 e 3:

$$\begin{aligned} x - y - z &= 2 \\ y + 3z &= 5 \\ 5z &= 10 \end{aligned} \qquad \left[\begin{array}{rrr|r} 1 & -1 & -1 & 2 \\ 0 & 1 & 3 & 5 \\ 0 & 0 & 5 & 10 \end{array}\right]$$

Esse é o mesmo sistema que resolvemos usando substituição de trás para a frente no Exemplo 5, em que descobrimos que a solução era [3, –1, 2]. Portanto, essa também é a solução do sistema dado neste exemplo. Por quê? Os cálculos que acabamos de ver mostram que *qualquer solução do sistema dado é também uma solução do sistema final*. No entanto, como os passos que acabamos de dar são *reversíveis*, poderíamos recuperar o sistema original, começando pelo sistema final. (Como?) Assim, *qualquer solução do sistema final é também uma solução do sistema dado*. Portanto, os sistemas são equivalentes (assim como são todos os sistemas obtidos nos passos intermediários que acabamos de ver). (Trabalhar com matrizes é o assunto da próxima seção.) Além disso, podemos trabalhar com matrizes tão bem quanto com equações, já que se trata apenas de recolocar as variáveis antes de fazer a substituição de trás para a frente. ◆

Observação: Calculadoras com recursos para operar com matrizes e sistemas computadorizados algébricos (CAS – *computer algebra system*) podem facilitar a resolução de sistemas de equações lineares, particularmente quando os sistemas são grandes ou têm coeficientes que não são "agradáveis", como freqüentemente ocorre nas aplicações encontradas na vida real. Como sempre, entretanto, você deve fazer tantos exemplos quanto puder com lápis e papel até que se sinta confortável com as técnicas. Mesmo quando uma calculadora ou um CAS for utilizado, pense sobre *como* você faria os cálculos manualmente antes de fazer qualquer coisa. Depois de obter uma resposta, não se esqueça de avaliar se ela é razoável.

Material com direitos autorais

Não se iluda pensando que a tecnologia sempre lhe dará a resposta de modo mais rápido ou mais fácil do que fazendo os cálculos à mão. Às vezes ela pode não fornecer nenhuma resposta! Erros de arredondamento associados à aritmética de ponto flutuante usada por calculadoras e computadores podem causar sérios problemas e conduzir a respostas completamente erradas. Veja Investigação: Mentiras que meu Computador me Contou para ter uma idéia do problema. (Você foi avisado!)

◆ EXERCÍCIOS 2.2 ◆

Nos Exercícios de 1 a 6, determine quais equações são lineares nas variáveis x, y e z. Se alguma equação não for linear, explique o motivo.

1. $x - \pi y + \sqrt[3]{5}z = 0$ **2.** $x^2 + y^2 + z^2 = 1$

3. $x^{-1} + 7y + z = \operatorname{sen}\left(\frac{\pi}{9}\right)$

4. $2x - xy - 5z = 0$ **5.** $3\cos x - 4y + z = \sqrt{3}$

6. $(\cos 3)x - 4y + z = \sqrt{3}$

Nos Exercícios de 7 a 10, encontre uma equação linear que tenha o mesmo conjunto solução que a equação dada (possivelmente com alguma restrição nas variáveis).

7. $2x + y = 7 - 3y$ **8.** $\frac{x^2 - y^2}{x - y} = 1$

9. $\frac{1}{x} + \frac{1}{y} = \frac{4}{xy}$ **10.** $\log_{10} x - \log_{10} y = 2$

Nos Exercícios de 11 a 14, encontre o conjunto solução de cada equação.

11. $3x - 6y = 0$ **12.** $2x_1 + 3x_2 = 5$

13. $x + 2y + 3z = 4$ **14.** $4x_1 + 3x_2 + 2x_3 = 1$

Nos Exercícios de 15 a 18, desenhe gráficos correspondentes aos sistemas lineares dados. Determine geometricamente se cada sistema tem uma única solução. Então resolva cada sistema algebricamente para confirmar sua resposta.

15. $\begin{aligned} x + y &= 0 \\ 2x + y &= 3 \end{aligned}$ **16.** $\begin{aligned} x - 2y &= 7 \\ 3x + y &= 7 \end{aligned}$

17. $\begin{aligned} 3x - 6y &= 3 \\ -x + 2y &= 1 \end{aligned}$ **18.** $\begin{aligned} 0{,}10x - 0{,}05y &= 0{,}20 \\ -0{,}06x + 0{,}03y &= -0{,}12 \end{aligned}$

Nos Exercícios de 19 a 24, resolva o sistema dado usando substituição.

19. $\begin{aligned} x - 2y &= 1 \\ y &= 3 \end{aligned}$ **20.** $\begin{aligned} 2u - 3v &= 5 \\ 2v &= 6 \end{aligned}$

21. $\begin{aligned} x - y + z &= 0 \\ 2y - z &= 1 \\ 3z &= -1 \end{aligned}$ **22.** $\begin{aligned} x_1 + 2x_2 + 3x_3 &= 0 \\ -5x_2 + 2x_3 &= 0 \\ 4x_3 &= 0 \end{aligned}$

23. $\begin{aligned} x_1 + x_2 - x_3 - x_4 &= 1 \\ x_2 + x_3 + x_4 &= 0 \\ x_3 - x_4 &= 0 \\ x_4 &= 1 \end{aligned}$ **24.** $\begin{aligned} x - 3y + z &= 5 \\ y - 2z &= -1 \end{aligned}$

Os sistemas dos Exercícios 25 e 26 exibem um padrão "triangular inferior" que torna fácil resolvê-los por substituição de frente para trás. (Encontraremos substituição de frente para trás de novo no Capítulo 3.) Resolva esses sistemas.

25. $\begin{aligned} x &= 2 \\ 2x + y &= -3 \\ -3x - 4y + z &= -10 \end{aligned}$ **26.** $\begin{aligned} x_1 &= -1 \\ -\tfrac{1}{2}x_1 + x_2 &= 5 \\ \tfrac{3}{2}x_1 + 2x_2 + x_3 &= 7 \end{aligned}$

Encontre as matrizes completas dos sistemas lineares dos Exercícios 27 a 30.

27. $\begin{aligned} x - y &= 0 \\ 2x + y &= 3 \end{aligned}$ **28.** $\begin{aligned} 2x_1 + 3x_2 - x_3 &= 1 \\ x_1 + x_3 &= 0 \\ -x_1 + 2x_2 - 2x_3 &= 0 \end{aligned}$

29. $\begin{aligned} x + 5y &= -1 \\ -x + y &= -5 \\ 2x + 4y &= 4 \end{aligned}$ **30.** $\begin{aligned} a - 2b + d &= 2 \\ -a + b - c - 3d &= 1 \end{aligned}$

Nos Exercícios 31 e 32, encontre um sistema de equações lineares que tenha a matriz dada como sua matriz completa.

31. $\left[\begin{array}{ccc|c} 0 & 1 & 1 & 1 \\ 1 & -1 & 0 & 1 \\ 2 & -1 & 1 & 1 \end{array}\right]$ **32.** $\left[\begin{array}{ccccc|c} 1 & -1 & 0 & 3 & 1 & 2 \\ 1 & 1 & 2 & 1 & -1 & 4 \\ 0 & 1 & 0 & 2 & 3 & 0 \end{array}\right]$

Nos Exercícios de 33 a 38, resolva os sistemas lineares dos exercícios indicados.

33. Exercício 27 **34.** Exercício 28

35. Exercício 29 **36.** Exercício 30

37. Exercício 31 **38.** Exercício 32

39. (a) Encontre um sistema de duas equações lineares nas variáveis x e y cujo conjunto solução seja dado pelas equações paramétricas $x = t$ e $y = 3 - 2t$.

Material com direitos autorais

(b) Encontre uma outra solução paramétrica para o sistema do item (a) na qual o parâmetro seja s e $y = s$.

40. (a) Encontre um sistema de duas equações lineares, nas variáveis x_1, x_2 e x_3 cujo conjunto solução seja dado pelas equações paramétricas $x_1 = t$, $x_2 = 1 + t$ e $x_3 = 2 - t$.

(b) Encontre uma outra solução paramétrica para o sistema do item (a) na qual o parâmetro seja s e $x_3 = s$.

Nos Exercícios de 41 a 44, os sistemas não são sistemas de equações lineares. Encontre substituições (mudanças de variáveis) que convertam cada sistema em um sistema linear e use esse sistema linear para ajudar a resolver o sistema dado.

41. $\begin{aligned} \frac{2}{x} + \frac{3}{y} &= 0 \\ \frac{3}{x} + \frac{4}{y} &= 1 \end{aligned}$

42. $\begin{aligned} x^2 + 2y^2 &= 6 \\ x^2 - y^2 &= 3 \end{aligned}$

43. $\begin{aligned} \operatorname{tg} x - 2 \operatorname{sen} y &= 2 \\ \operatorname{tg} x - \operatorname{sen} y + \cos z &= 2 \\ \operatorname{sen} y - \cos z &= -1 \end{aligned}$

44. $\begin{aligned} -2^a + 2(3^b) &= 1 \\ 3(2^a) - 4(3^b) &= 1 \end{aligned}$

Material com direitos autorais

Investigação

Mentiras que meu Computador me Contou

Computadores e calculadoras armazenam números reais em uma *representação em ponto flutuante*. Por exemplo, o número 2001 é armazenado como $0{,}2001 \times 10^4$, e o número $-0{,}00063$ é armazenado como $-0{,}63 \times 10^{-3}$. Em geral, uma representação em ponto flutuante de um número real é $\pm M \times 10^k$, onde k é um inteiro e a *mantissa* M é um número real (decimal) que satisfaz $0{,}1 \leq M < 1$.

O número máximo de casas decimais que podem ser armazenadas na mantissa depende do computador, da calculadora ou do sistema computacional algébrico. Se o número máximo de casas decimais que podem ser armazenadas for d, dizemos que há *d algarismos significativos*. Muitas calculadoras armazenam oito ou 12 algarismos significativos. Computadores podem armazenar mais, mas ainda estão sujeitos a um limite. Algarismos não armazenados são omitidos (neste caso, dizemos que o número foi *truncado*) ou usados para *arredondamento*, de modo que o número fique com d algarismos significativos.

Por exemplo, $\pi \approx 3{,}141592654$, e sua representação em ponto flutuante é $0{,}3141592654 \times 10^1$. Em um computador que reduz a cinco algarismos significativos por truncamento, o número π é armazenado como $0{,}31415 \times 10^1$ (mostrado no visor como 3,1415); um computador que arredonda para cinco algarismos significativos armazenaria π como $0{,}31416 \times 10^1$ (mostrado no visor como 3,1416). Quando o algarismo desprezado é um 5, o arrendondamento é feito de modo que o último algarismo que permanece seja par. Assim, arredondando para dois algarismos significativos, o número 0,735 torna-se 0,74, enquanto 0,725 torna-se 0,72.

Toda vez que um truncamento ou um arredondamento ocorre, um *erro de arredondamento* é iniciado, podendo ter um efeito dramático nos cálculos. Quanto mais operações forem realizadas, mais o erro se acumulará. Às vezes, infelizmente, não há o que possamos fazer a esse respeito. Esta investigação ilustra esse fenômeno com um sistema de equações lineares bastante simples.

1. Resolva o seguinte sistema de equações lineares *exatamente* (isto é, trabalhe com números fracionários durante os cálculos).

$$\begin{aligned} x + \quad y &= 0 \\ x + \tfrac{801}{800}y &= 1 \end{aligned}$$

2. Na forma decimal, $\frac{801}{800} = 1{,}00125$. Portanto, se arredondarmos para cinco algarismos significativos, o sistema passará a ser

$$\begin{aligned} x + \qquad y &= 0 \\ x + 1{,}0012y &= 1 \end{aligned}$$

Usando sua calculadora ou CAS, resolva esse sistema, arredondando o resultado de cada cálculo para cinco algarismos significativos.

3. Resolva o sistema mais duas vezes, arredondando primeiramente para quatro algarismos significativos e depois para três algarismos significativos. O que acontece?

4. Sem dúvida, um pequeno erro de arredondamento (menor ou igual a 0,00125) pode levar a erros bastante grandes no resultado. Explique o motivo geometricamente. (Pense nos gráficos dos vários sistemas lineares que você resolveu nos problemas de 1 a 3.)

Material com direitos autorais

Sistemas como esse que você acabou de ver são chamados ***malcondicionados***. Eles são extremamente sensíveis para arredondamentos e não há muito que possamos fazer. Encontraremos sistemas malcondicionados nos Capítulos 3 e 7. Aqui está outro exemplo para praticar:

$$4{,}552x + 7{,}083y = 1{,}931$$

$$1{,}731x + 2{,}693y = 2{,}001$$

Brinque com vários números de algarismos significativos para ver o que acontece, começando com oito algarismos significativos (se puder).

Material com direitos autorais

2.3 Métodos Diretos de Resolução de Sistemas Lineares

Nesta seção, veremos um procedimento geral e sistemático de resolução de sistema de equações lineares. Esse procedimento é baseado na idéia de reduzir a matriz completa do sistema dado a uma forma que possa então ser resolvida por substituição de trás para a frente. O método é *direto* no sentido de que leva diretamente à solução (se existir) em um número finito de passos. Na Seção 2.5, consideraremos alguns métodos *indiretos* que funcionam de um modo completamente diferente.

As Matrizes e a Forma Escalonada

Há duas matrizes importantes associadas a um sistema linear. A ***matriz dos coeficientes*** contém os coeficientes das variáveis, e a ***matriz completa*** (que já vimos) é a matriz dos coeficientes acrescentada de uma coluna extra que contém os termos constantes.

Para o sistema

$$\begin{aligned} 2x + y - z &= 3 \\ x + 5z &= 1 \\ -x + 3y - 2z &= 0 \end{aligned}$$

a matriz dos coeficientes é

$$\begin{bmatrix} 2 & 1 & -1 \\ 1 & 0 & 5 \\ -1 & 3 & -2 \end{bmatrix}$$

e a matriz completa é

$$\left[\begin{array}{rrr|r} 2 & 1 & -1 & 3 \\ 1 & 0 & 5 & 1 \\ -1 & 3 & -2 & 0 \end{array}\right]$$

A palavra *escalonar* vem da palavra latina *scala*, que significa "escada" ou "degraus". Escalonar uma matriz significa dar a ela a forma de escada.

Note que, se uma variável estiver faltando (como a variável y na segunda equação), seu coeficiente 0 será colocado na matriz na posição correspondente. Se denotarmos a matriz dos coeficientes de um sistema linear por A e o vetor coluna dos termos constantes por $\mathbf{b}$, a forma da matriz completa será $[A \mid \mathbf{b}]$.

Na resolução de um sistema linear, nem sempre será possível reduzir a matriz dos coeficientes à forma triangular, como fizemos no Exemplo 6 da Seção 2.2. Entretanto, sempre podemos conseguir um formato de escada nos elementos não nulos da matriz final.

Definição Uma matriz está na *forma escalonada por linhas* quando satisfaz as seguintes propriedades:

1. Todas as linhas que consistem inteiramente em zeros estão na parte inferior da matriz.
2. Em cada linha não nula, o primeiro elemento não nulo (chamado de *elemento líder*) está em uma coluna à esquerda de qualquer outro elemento líder abaixo dele.

Material com direitos autorais

Note que essas propriedades garantem que os elementos líderes fiquem posicionados formando uma escada. Em particular, em qualquer coluna que contenha um elemento líder, todos os elementos abaixo dele são nulos, como ilustram os exemplos a seguir.

EXEMPLO 1 As seguintes matrizes estão na forma escalonada por linhas:

$$\begin{bmatrix} 2 & 4 & 1 \\ 0 & -1 & 2 \\ 0 & 0 & 0 \end{bmatrix} \quad \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 5 \\ 0 & 0 & 4 \end{bmatrix} \quad \begin{bmatrix} 1 & 1 & 2 & 1 \\ 0 & 0 & 1 & 3 \\ 0 & 0 & 0 & 0 \end{bmatrix} \quad \begin{bmatrix} 0 & 2 & 0 & 1 & -1 & 3 \\ 0 & 0 & -1 & 1 & 2 & 2 \\ 0 & 0 & 0 & 0 & 4 & 0 \\ 0 & 0 & 0 & 0 & 0 & 5 \end{bmatrix}$$ ◆

Se uma matriz escalonada por linhas for a matriz completa de um sistema linear, será fácil resolver o sistema usando apenas o método de substituição.

EXEMPLO 2 Supondo que cada uma das matrizes do Exemplo 1 seja uma matriz completa, escreva os sistemas de equações lineares correspondentes e resolva-os.

SOLUÇÃO: Vamos primeiramente lembrar que a última coluna de uma matriz completa é o vetor dos termos constantes. A primeira matriz então corresponde ao sistema

$$\begin{aligned} 2x_1 + 4x_2 &= 1 \\ -x_2 &= 2 \end{aligned}$$

(Note que desprezamos a última equação $0 = 0$, ou $0x_1 + 0x_2 = 0$, que é claramente satisfeita por quaisquer valores de x_1 e de x_2.) A substituição fornece $x_2 = -2$ e então $2x_1 = 1 - 4(-2) = 9$, logo, $x_1 = \frac{9}{2}$. A solução é $[\frac{9}{2}, -2]$.

A segunda matriz corresponde ao seguinte sistema:

$$\begin{aligned} x_1 \qquad &= 1 \\ x_2 \quad &= 5 \\ 0 &= 4 \end{aligned}$$

A última equação representa $0x_1 + 0x_2 = 4$, que claramente não tem soluções. Assim, o sistema não tem soluções. Analogamente, o sistema que corresponde à quarta matriz não tem soluções. Para o sistema que corresponde à terceira matriz, temos:

$$\begin{aligned} x_1 + x_2 + 2x_3 &= 1 \\ x_3 &= 3 \end{aligned}$$

logo, $x_1 = 1 - 2(3) - x_2 = -5 - x_2$. Há infinitas soluções, já que podemos associar x_2 a qualquer valor t e encontrar a solução paramétrica $[-5 - t, t, 3]$. ◆

Operações Elementares com as Linhas

Vamos agora descrever o procedimento pelo qual qualquer matriz pode ser reduzida a uma matriz na forma escalonada por linhas. As operações permitidas, chamadas ***operações elementares com as linhas***, correspondem às operações que podem ser realizadas em um sistema de equações lineares para transformá-lo em um sistema equivalente.

Material com direitos autorais

Definição As seguintes ***operações elementares com as linhas*** podem ser realizadas em uma matriz:

1. Trocar duas linhas.
2. Multiplicar uma linha por uma constante não nula.
3. Somar um múltiplo de uma linha com outra linha.

Observe que dividir uma linha por uma constante não nula é uma propriedade que está incluída nessa definição, já que, por exemplo, dividir uma linha por 2 é o mesmo que multiplicá-la por $\frac{1}{2}$. Analogamente, subtrair de uma linha um múltiplo de outra linha é o mesmo que somar a uma linha um múltiplo negativo de outra.

Usaremos a seguinte notação para as três operações elementares com linhas:

1. $L_i \leftrightarrow L_j$ significa trocar as linhas i e j.
2. kL_i significa multiplicar a linha i pelo número real k.
3. $L_i + kL_j$ significa somar k vezes a linha j à linha i (e trocar a linha i pelo resultado).

O processo de aplicar operações elementares com linhas para transformar uma matriz em uma matriz escalonada é chamado ***escalonamento***.

EXEMPLO 3 Reduza a seguinte matriz à forma escalonada:

$$\begin{bmatrix} 1 & 2 & -4 & -4 & 5 \\ 2 & 4 & 0 & 0 & 2 \\ 2 & 3 & 2 & 1 & 5 \\ -1 & 1 & 3 & 6 & 5 \end{bmatrix}$$

SOLUÇÃO: Trabalhamos coluna por coluna, da esquerda para a direita e de cima para baixo. A estratégia é criar um elemento líder em uma coluna e usá-lo para criar zeros sob ele. O elemento escolhido para ser o elemento líder é chamado ***pivô***, e essa fase do processo é chamada ***pivoteamento***. Embora não seja estritamente necessário, às vezes é conveniente usar a segunda operação elementar para transformar o elemento líder em 1.

Começamos por anular os elementos da primeira coluna, abaixo do líder 1 na primeira linha:

$$\begin{bmatrix} 1 & 2 & -4 & -4 & 5 \\ 2 & 4 & 0 & 0 & 2 \\ 2 & 3 & 2 & 1 & 5 \\ -1 & 1 & 3 & 6 & 5 \end{bmatrix} \xrightarrow[L_4+L_1]{\substack{L_2-2L_1 \\ L_3-2L_1}} \begin{bmatrix} 1 & 2 & -4 & -4 & 5 \\ 0 & 0 & 8 & 8 & -8 \\ 0 & -1 & 10 & 9 & -5 \\ 0 & 3 & -1 & 2 & 10 \end{bmatrix}$$

A primeira coluna agora está como queremos; o próximo passo é criar um elemento líder na segunda linha, com o objetivo de chegar à forma escalonada. Neste caso, fazemos isso trocando as linhas. (Poderíamos também somar a linha 3 ou 4 à linha 2.)

$$\xrightarrow{L_2 \leftrightarrow L_3} \begin{bmatrix} 1 & 2 & -4 & -4 & 5 \\ 0 & -1 & 10 & 9 & -5 \\ 0 & 0 & 8 & 8 & -8 \\ 0 & 3 & -1 & 2 & 10 \end{bmatrix}$$

Desta vez o pivô é –1. Criamos zeros no restante da segunda coluna usando o elemento líder –1 na linha 2:

$$\xrightarrow{L_4+3L_2} \begin{bmatrix} 1 & 2 & -4 & -4 & 5 \\ 0 & -1 & 10 & 9 & -5 \\ 0 & 0 & 8 & 8 & -8 \\ 0 & 0 & 29 & 29 & -5 \end{bmatrix}$$

Material com direitos autorais

A coluna 2 agora está pronta. Notando que já temos um elemento líder na coluna 3, nós o escolhemos como pivô e começamos a tarefa de colocar zeros sob ele. Isso ficará mais fácil se primeiramente dividirmos a linha 3 por 8:

$$\xrightarrow{\frac{1}{8}L_3}\begin{bmatrix} 1 & 2 & -4 & -4 & 5 \\ 0 & -1 & 10 & 9 & -5 \\ 0 & 0 & 1 & 1 & -1 \\ 0 & 0 & 29 & 29 & -5 \end{bmatrix}$$

Agora usamos o elemento líder 1 na coluna 3 para criar zeros sob ele:

$$\xrightarrow{L_4-29L_3}\begin{bmatrix} 1 & 2 & -4 & -4 & 5 \\ 0 & -1 & 10 & 9 & -5 \\ 0 & 0 & 1 & 1 & -1 \\ 0 & 0 & 0 & 0 & 24 \end{bmatrix}$$

Com esse passo final, reduzimos a matriz dada à forma escalonada. ◆

Observações:

- A forma escalonada da matriz não é única. (Encontre uma forma escalonada diferente para a matriz do Exemplo 3.)
- O elemento líder de cada linha é usado para criar zeros abaixo dele.
- Os pivôs não são necessariamente os elementos que estavam originalmente nas posições ocupadas pelos elementos líderes. No Exemplo 3, os pivôs eram 1, –1, 8 e 24. A matriz original tinha 1, 4, 2 e 5 nessas posições.
- Uma vez que tenhamos realizado o pivoteamento e anulado os elementos sob o elemento líder em uma coluna, aquela coluna não muda mais. Em outras palavras, a forma escalonada por linhas é construída da esquerda para a direita e de cima para baixo.

Operações elementares com linhas são reversíveis — isto é, podem ser "desfeitas". Assim, se uma operação elementar sobre as linhas converte A em B, existe também uma operação elementar sobre as linhas que converte B em A. (Veja os Exercícios 15 e 16.)

Definição As matrizes A e B serão ***linha-equivalentes*** se existir uma seqüência de operações elementares com as linhas que converta A em B.

As matrizes do Exemplo 3,

$$\begin{bmatrix} 1 & 2 & -4 & -4 & 5 \\ 2 & 4 & 0 & 0 & 2 \\ 2 & 3 & 2 & 1 & 5 \\ -1 & 1 & 3 & 6 & 5 \end{bmatrix} \quad \text{e} \quad \begin{bmatrix} 1 & 2 & -4 & -4 & 5 \\ 0 & -1 & 10 & 9 & -5 \\ 0 & 0 & 1 & 1 & -1 \\ 0 & 0 & 0 & 0 & 24 \end{bmatrix}$$

são linha-equivalentes. Entretanto, em geral, como podemos dizer se duas matrizes são linha-equivalentes?

Material com direitos autorais

◆ **TEOREMA 1**

As matrizes A e B são linha-equivalentes se, e somente se, puderem ser reduzidas à mesma forma escalonada por linhas.

DEMONSTRAÇÃO: Se A e B são linha-equivalentes, operações adicionais com linhas reduzirão B (e conseqüentemente A) à (mesma) forma escalonada por linhas.

Reciprocamente, se A e B têm a mesma forma escalonada por linhas R, então, por meio de operações elementares por linhas, podemos converter A em R e B em R. Revertendo a última seqüência de operações, podemos converter R em B, e, portanto, a seqüência $A \to R \to B$ produz o efeito desejado. ◆

Observação: Na prática, é mais fácil usar o Teorema 1 se R for a forma escalonada *reduzida* de A e de B, como será definido na *página 72*. Veja os Exercícios 17 e 18.

O Método de Eliminação de Gauss

Quando uma redução por linhas é aplicada à matriz completa de um sistema de equações lineares, criamos um sistema equivalente que pode ser resolvido por substituição de trás para a frente. O processo inteiro é conhecido como ***método de eliminação de Gauss***, ou ***método de eliminação gaussiana***.

◆ **O Método de Eliminação de Gauss**

1. Escreva a matriz completa do sistema de equações lineares.
2. Use operações elementares com as linhas para reduzir a matriz completa à forma escalonada por linhas.
3. Usando substituição de trás para a frente, resolva o sistema equivalente que corresponde à matriz linha-reduzida.

EXEMPLO 4 Resolva o sistema

$$\begin{aligned} x_1 + x_2 + x_3 &= 3 \\ 2x_1 + 3x_2 + x_3 &= 5 \\ x_1 - x_2 - 2x_3 &= -5 \end{aligned}$$

SOLUÇÃO: A matriz completa é

$$\left[\begin{array}{ccc|c} 1 & 1 & 1 & 3 \\ 2 & 3 & 1 & 5 \\ 1 & -1 & -2 & -5 \end{array}\right]$$

Continuamos o processo reduzindo essa matriz à forma escalonada por linhas:

$$\left[\begin{array}{ccc|c} 1 & 1 & 1 & 3 \\ 2 & 3 & 1 & 5 \\ 1 & -1 & -2 & -5 \end{array}\right] \xrightarrow[L_3 - L_1]{L_2 - 2L_1} \left[\begin{array}{ccc|c} 1 & 1 & 1 & 3 \\ 0 & 1 & -1 & -1 \\ 0 & -2 & -3 & -8 \end{array}\right] \xrightarrow{L_3 + 2L_2} \left[\begin{array}{ccc|c} 1 & 1 & 1 & 3 \\ 0 & 1 & -1 & -1 \\ 0 & 0 & -5 & -10 \end{array}\right]$$

Material com direitos autorais

Carl Friedrich Gauss (1777–1855) é considerado um dos três maiores matemáticos de todos os tempos, juntamente com Arquimedes e Newton. É freqüentemente chamado de "príncipe dos matemáticos", apelido que ele merece. Criança prodígio, conta-se que Gauss conseguia fazer aritmética antes de conseguir falar. Com três anos de idade, ele corrigiu um erro nos cálculos feitos por seu pai para a folha de pagamentos da companhia, e, quando jovem estudante, deduziu a fórmula $n(n+1)/2$ para a soma dos n primeiros números naturais. Com 19 anos, provou que um polígono de 17 lados* poderia ser construído com o uso de apenas uma régua e um compasso; aos 21, provou, em sua dissertação de doutorado, que todo polinômio de grau n com coeficientes reais ou complexos tem exatamente n raízes, contando suas multiplicidades – o Teorema Fundamental da Álgebra.

A publicação de Gauss *Disquisitiones arithmeticae*, de 1801, é em geral considerada o fundamento da teoria moderna dos números, mas ele fez contribuições em praticamente todos os ramos da matemática, bem como à estatística, física, astronomia e agrimensura. Gauss não publicou todos os seus resultados, possivelmente por ser muito crítico a respeito de seu próprio trabalho. Ele também não gostava de dar aulas e freqüentemente era crítico de outros matemáticos, talvez por ter descoberto — mas não publicado — os resultados antes deles.

O método de eliminação de Gauss era conhecido pelos chineses no terceiro século a.C., mas carrega o nome de Gauss por causa de sua redescoberta em um artigo no qual ele resolveu um sistema de equações lineares para descrever a órbita de um asteróide.

O sistema correspondente é agora

$$\begin{aligned} x_1 + x_2 + x_3 &= 3 \\ x_2 - x_3 &= -1 \\ -5x_3 &= -10 \end{aligned}$$

A substituição de trás para a frente nos dá $x_3 = 2$, $x_2 = 1$, $x_1 = 0$, e, portanto, a solução pode ser escrita na forma vetorial como

$$\begin{bmatrix} 0 \\ 1 \\ 2 \end{bmatrix}$$

De agora em diante, escreveremos soluções vetoriais de sistemas de equações lineares como vetores-coluna. (O motivo disso ficará claro no Capítulo 3.)

EXEMPLO 5 Resolva o sistema

$$\begin{aligned} w - x - y + 2z &= 1 \\ 2w - 2x - y + 3z &= 3 \\ -w + x - y \qquad &= -3 \end{aligned}$$

SOLUÇÃO: A matriz completa é

$$\left[\begin{array}{rrrr|r} 1 & -1 & -1 & 2 & 1 \\ 2 & -2 & -1 & 3 & 3 \\ -1 & 1 & -1 & 0 & -3 \end{array}\right]$$

que pode ser reduzida por linhas, como a seguir:

$$\left[\begin{array}{rrrr|r} 1 & -1 & -1 & 2 & 1 \\ 2 & -2 & -1 & 3 & 3 \\ -1 & 1 & -1 & 0 & -3 \end{array}\right] \xrightarrow[R_3+R_1]{R_2-2R_1} \left[\begin{array}{rrrr|r} 1 & -1 & -1 & 2 & 1 \\ 0 & 0 & 1 & -1 & 1 \\ 0 & 0 & -2 & 2 & -2 \end{array}\right] \xrightarrow{R_3+2R_2} \left[\begin{array}{rrrr|r} 1 & -1 & -1 & 2 & 1 \\ 0 & 0 & 1 & -1 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right]$$

O sistema associado agora é

$$\begin{aligned} w - x - y + 2z &= 1 \\ y - z &= 1 \end{aligned}$$

que tem infinitas soluções. Há mais de uma maneira de atribuir parâmetros, mas continuaremos usando a substituição de trás para a frente, escrevendo as variáveis correspondentes aos elementos líderes (as ***variáveis dependentes***) em termos das outras variáveis (as ***variáveis livres***).

* N.T.: No original, o autor escreveu apenas "polígono de 17 lados", mas "polígono regular de 17 lados" seria mais apropriado.

Material com direitos autorais

Neste caso, as variáveis dependentes são w e y, e as variáveis livres, x e z. Então, $y = 1 + z$, e disso obtemos

$$\begin{aligned} w &= 1 + x + y - 2z \\ &= 1 + x + (1 + z) - 2z \\ &= 2 + x - z \end{aligned}$$

Se atribuirmos parâmetros $x = s$ e $z = t$, a solução pode ser escrita na forma vetorial como

$$\begin{bmatrix} w \\ x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 2 + s - t \\ s \\ 1 + t \\ t \end{bmatrix} = \begin{bmatrix} 2 \\ 0 \\ 1 \\ 0 \end{bmatrix} + s\begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \end{bmatrix} + t\begin{bmatrix} -1 \\ 0 \\ 1 \\ 1 \end{bmatrix}$$ ◆

O Exemplo 5 destaca uma propriedade muito importante: em um sistema possível, as variáveis livres são justamente as que não são variáveis dependentes. Como o número de variáveis dependentes é igual ao número de linhas não nulas na forma escalonada da matriz dos coeficientes, podemos predizer o número de variáveis livres (parâmetros) antes de acharmos a solução explícita usando substituição de trás para a frente. No Capítulo 3, provaremos que, embora a forma escalonada da matriz não seja única, o número de linhas não nulas é o mesmo em *qualquer* forma escalonada de uma matriz dada. Portanto, faz sentido atribuir um nome a esse número.

Definição O ***posto*** de uma matriz é o número de linhas não nulas de qualquer uma de suas formas escalonadas por linhas.

Denotaremos o posto de uma matriz A por posto(A). No Exemplo 4, o posto da matriz dos coeficientes é 3; no Exemplo 5, o posto da matriz dos coeficientes é 2. As observações que acabamos de fazer justificam o teorema a seguir, que demonstraremos com mais generalidade nos Capítulos 3 e 6.

◆ TEOREMA 2 O Teorema do Posto

Seja A a matriz dos coeficientes de um sistema de equações lineares com n variáveis. Se o sistema for possível, então o

$$\text{número de variáveis livres} = n - \text{posto}(A)$$

Assim, no Exemplo 4, temos $3 - 3 = 0$ variáveis livres (em outras palavras, uma *única* solução), e, no Exemplo 5, temos $4 - 2 = 2$ variáveis livres, como havíamos encontrado.

EXEMPLO 6 Resolva o sistema

$$\begin{aligned} x_1 - x_2 + 2x_3 &= 3 \\ x_1 + 2x_2 - x_3 &= -3 \\ 2x_2 - 2x_3 &= 1 \end{aligned}$$

SOLUÇÃO: Quando reduzimos a matriz completa à forma escalonada por linhas, obtemos:

$$\left[\begin{array}{rrr|r} 1 & -1 & 2 & 3 \\ 1 & 2 & -1 & -3 \\ 0 & 2 & -2 & 1 \end{array}\right] \xrightarrow{R_2 - R_1} \left[\begin{array}{rrr|r} 1 & -1 & 2 & 3 \\ 0 & 3 & -3 & -6 \\ 0 & 2 & -2 & 1 \end{array}\right]$$

Material com direitos autorais

$$\xrightarrow{\frac{1}{3}R_2} \left[\begin{array}{rrr|r} 1 & -1 & 2 & 3 \\ 0 & 1 & -1 & -2 \\ 0 & 2 & -2 & 1 \end{array}\right]$$

$$\xrightarrow{R_3-2R_2} \left[\begin{array}{rrr|r} 1 & -1 & 2 & 3 \\ 0 & 1 & -1 & -2 \\ 0 & 0 & 0 & 5 \end{array}\right]$$

levando à equação impossível 0 = 5. (Poderíamos também ter feito $L_3 - \frac{2}{3}L_2$ como segunda operação elementar, o que nos teria levado à mesma contradição, mas a uma forma escalonada diferente.) Assim, o sistema não tem solução — ele é impossível. ◆

O Método de Eliminação de Gauss-Jordan

Uma modificação do método de eliminação de Gauss simplifica bastante a fase de substituição de trás para a frente e é particularmente útil quando os cálculos estão sendo feitos à mão em um sistema com infinitas soluções. Essa variante, conhecida como ***método de eliminação de Gauss-Jordan***, baseia-se em reduzir ainda mais a matriz completa.

Definição Uma matriz está na ***forma escalonada reduzida*** (por linhas) se ela satisfaz às seguintes propriedades:

1. Quaisquer linhas que consistem inteiramente de zeros estão na parte inferior da matriz.
2. O elemento líder em cada linha não nula é igual a 1 (chamado ***1 líder***).
3. Cada coluna que contém um 1 líder tem zeros em todas as outras posições.

A seguinte matriz está na forma escalonada reduzida:

$$\begin{bmatrix} 1 & 2 & 0 & 0 & -3 & 1 & 0 \\ 0 & 0 & 1 & 0 & 4 & -1 & 0 \\ 0 & 0 & 0 & 1 & 3 & -2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Para matrizes 2 × 2, as formas escalonadas reduzidas possíveis são:

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad \begin{bmatrix} 1 & * \\ 0 & 0 \end{bmatrix}, \quad \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \quad \text{e} \quad \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$

onde * pode ser qualquer número.

Wilhelm Jordan (1842–1899) era um professor alemão de geodésia. Sua contribuição à resolução de sistemas lineares foi um método sistemático de substituição de trás para a frente estreitamente relacionado com o método descrito aqui.

É claro que, depois de uma matriz ter sido reduzida à forma escalonada, mais operações elementares com as linhas irão levá-la à forma escalonada reduzida. O que não é claro (embora a intuição possa sugerir) é que, diferentemente do que ocorria com a forma escalonada por linhas, a forma escalonada reduzida de uma matriz é *única*.[1]

[1] Para uma pequena prova desse fato, veja o artigo "The reduced row echelon form of a matrix is unique: A simple proof", de Thomas Yuster, na edição de março de 1984 da revista *Mathematics magazine* (vol. 57, n. 2, p. 93-94).

Material com direitos autorais

No método de eliminação de Gauss-Jordan, procedemos como no método de eliminação de Gauss, mas reduzimos ainda mais a matriz completa até a forma escalonada *reduzida*.

◆ Método de eliminação de Gauss-Jordan

1. Escreva a matriz completa do sistema de equações lineares.
2. Use operações elementares com linhas para reduzir a matriz completa à forma escalonada reduzida.
3. Se o sistema resultante for possível, resolva-o para as variáveis dependentes em termos de quaisquer variáveis livres que tenham sobrado.

EXEMPLO 7 Resolva o sistema do Exemplo 5 pelo método de eliminação de Gauss-Jordan.

SOLUÇÃO: A redução é feita como fizemos no Exemplo 5, até chegarmos à forma escalonada:

$$\left[\begin{array}{cccc|c} 1 & -1 & -1 & 2 & 1 \\ 0 & 0 & 1 & -1 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right]$$

Agora, devemos criar um 0 acima do 1 líder da segunda linha, terceira coluna. Fazemos isso somando a linha 2 à linha 1, obtendo:

$$\left[\begin{array}{cccc|c} 1 & -1 & 0 & 1 & 2 \\ 0 & 0 & 1 & -1 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right]$$

O sistema foi então reduzido a

$$\begin{aligned} w - x \quad + z &= 2 \\ y - z &= 1 \end{aligned}$$

Agora é bem mais fácil resolver o sistema em função das variáveis dependentes:

$$w = 2 + x - z \quad \text{e} \quad y = 1 + z$$

Se atribuirmos parâmetros $x = s$ e $z = t$, como anteriormente, a solução pode ser escrita na forma vetorial da seguinte maneira:

$$\begin{bmatrix} w \\ x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 2 + s - t \\ s \\ 1 + t \\ t \end{bmatrix}$$ ◆

Observação: • Do ponto de vista computacional, é mais eficiente (no sentido de requerer menos cálculos) primeiro reduzir a matriz à forma escalonada por linhas e, depois, trabalhando *da direita para a esquerda*, transformar cada elemento líder em 1 e criar zeros acima desses 1 líderes. Entretanto, para cálculos à mão, você achará mais fácil trabalhar da esquerda para a direita e criar o 1 líder e os zeros em suas colunas à medida que for trabalhando.

Material com direitos autorais

Retornemos à geometria que nos trouxe até este ponto. Assim como equações lineares* com duas variáveis correspondem a retas em $\mathbb{R}^2$, equações lineares em três variáveis correspondem a planos em $\mathbb{R}^3$. De fato, muitas questões sobre retas e planos podem ser respondidas através da resolução de um sistema linear apropriado.

EXEMPLO 8 Encontre a reta interseção dos planos $x + 2y - z = 3$ e $2x + 3y + z = 1$.

SOLUÇÃO: Observe primeiramente que *haverá* uma reta interseção, já que os vetores normais aos planos — $[1, 2, -1]$ e $[2, 3, 1]$ — não são paralelos. Os pontos que estão na interseção dos planos correspondem aos pontos do conjunto solução do sistema

$$\begin{aligned} x + 2y - z &= 3 \\ 2x + 3y + z &= 1 \end{aligned}$$

O método de eliminação de Gauss-Jordan, aplicado à matriz completa, fornece:

$$\left[\begin{array}{rrr|r} 1 & 2 & -1 & 3 \\ 2 & 3 & 1 & 1 \end{array}\right] \xrightarrow{L_2 - 2L_1} \left[\begin{array}{rrr|r} 1 & 2 & -1 & 3 \\ 0 & -1 & 3 & -5 \end{array}\right]$$

$$\xrightarrow[-L_2]{L_1 + 2L_2} \left[\begin{array}{rrr|r} 1 & 0 & 5 & -7 \\ 0 & 1 & -3 & 5 \end{array}\right]$$

Colocando novamente as variáveis, temos

$$\begin{aligned} x \qquad + 5z &= -7 \\ y - 3z &= 5 \end{aligned}$$

Igualamos a variável livre z a um parâmetro t e então obtemos as equações paramétricas da reta interseção dos dois planos:

$$\begin{aligned} x &= -7 - 5t \\ y &= 5 + 3t \\ z &= t \end{aligned}$$

Na forma vetorial, a equação é

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} -7 \\ 5 \\ 0 \end{bmatrix} + t\begin{bmatrix} -5 \\ 3 \\ 1 \end{bmatrix}$$

Veja a Figura 1.

* N.T.: No original está escrito "Assim como sistemas de equações lineares...", mas não é correto: o autor certamente quis dizer "equações", e não "sistemas de equações".

Material com direitos autorais

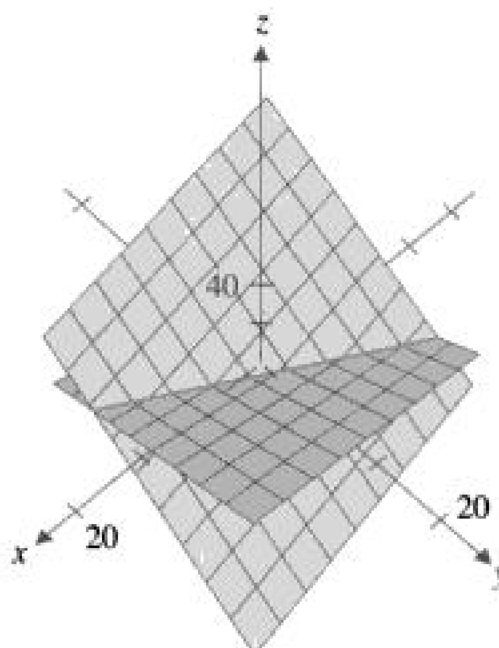


Figura 1 A interseção de dois planos ◆

EXEMPLO 9 Sejam $\mathbf{p} = \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}$, $\mathbf{q} = \begin{bmatrix} 0 \\ 2 \\ 1 \end{bmatrix}$, $\mathbf{u} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$, e $\mathbf{v} = \begin{bmatrix} 3 \\ -1 \\ -1 \end{bmatrix}$. Determine se as retas $\mathbf{x} = \mathbf{p} + t\mathbf{u}$ e $\mathbf{x} = \mathbf{q} + t\mathbf{v}$ se interceptam e, em caso afirmativo, encontre o ponto de interseção.

SOLUÇÃO: Precisamos ser cuidadosos aqui. Embora a letra t tenha sido usada como parâmetro nas equações das duas retas, as retas são independentes e, portanto, seus parâmetros também o são. Vamos usar outro parâmetro — por exemplo, s — para a primeira reta; sua equação torna-se $\mathbf{x} = \mathbf{p} + s\mathbf{u}$. Se as retas se interceptarem, precisaremos encontrar um $\mathbf{x} = \begin{bmatrix} x \\ y \\ z \end{bmatrix}$ que satisfaça ambas as equações simultaneamente — isto é, $\mathbf{x} = \mathbf{p} + s\mathbf{u} = \mathbf{q} + t\mathbf{v}$ ou $s\mathbf{u} - t\mathbf{v} = \mathbf{q} - \mathbf{p}$.

Substituindo **p**, **q**, **u** e **v** dados, obtemos as equações

$$\begin{aligned} s - 3t &= -1 \\ s + t &= 2 \\ s + t &= 2 \end{aligned}$$

cuja solução, fácil de encontrar, é $s = \frac{5}{4}$ e $t = \frac{3}{4}$. O ponto de interseção é, portanto,

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix} + \frac{5}{4}\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} \frac{9}{4} \\ \frac{5}{4} \\ \frac{1}{4} \end{bmatrix}$$

Veja a Figura 2. (Verifique que a substituição $t = \frac{3}{4}$ na outra equação dá o mesmo ponto.) ◆

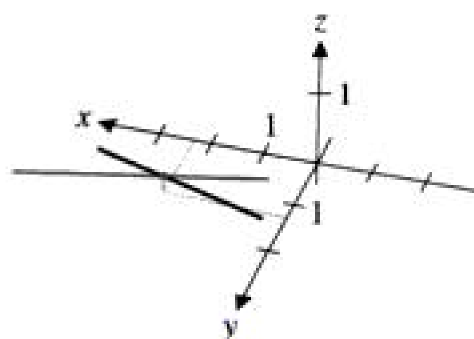


Figura 2 Duas retas concorrentes*

* N.T.: *Retas concorrentes* são retas que se interceptam.

Material com direitos autorais

Observação: • Em $\mathbb{R}^3$, duas retas distintas ou podem se interceptar em um ponto, ou ser paralelas, ou podem estar em uma posição relativa diferente das anteriores. Retas não paralelas que não se interceptam são chamadas *retas reversas*.

Sistemas Homogêneos

Vimos que todo sistema de equações lineares ou não tem nenhuma solução, ou tem uma única solução, ou tem infinitas soluções. Porém, existe um tipo de sistema linear que sempre tem pelo menos uma solução.

Definição Um sistema de equações lineares é chamado ***homogêneo*** se o termo constante em cada equação é igual a zero.

Em outras palavras, um sistema homogêneo tem uma matriz completa da forma $[A \mid \mathbf{0}]$. O seguinte sistema é homogêneo:

$$\begin{aligned} 2x + 3y - z &= 0 \\ -x + 5y + 2z &= 0 \end{aligned}$$

Como um sistema homogêneo não pode não ter solução (perdoe a dupla negação!), ele terá ou uma única solução ou infinitas soluções. O próximo teorema diz que o último caso *ocorrerá* sempre que o número de variáveis for maior que o número de equações.

◆ **TEOREMA 3**

Se $[A \mid \mathbf{0}]$ for a matriz completa de um sistema homogêneo de m equações lineares com n variáveis, onde $m < n$, então o sistema terá infinitas soluções.

DEMONSTRAÇÃO: Como o sistema admite pelo menos a solução nula, ele é possível. Além disso, posto$(A) \le m$ (por quê?). Pelo Teorema do Posto, temos:

$$\text{número de variáveis livres} = n - \text{posto}(A) \ge n - m > 0.$$

Existe, portanto, pelo menos uma variável livre e, conseqüentemente, infinitas soluções. ◆

Nota: O Teorema 3 não diz nada sobre o caso $m \ge n$. O Exercício 42 pede a você exemplos que mostrem que, nessa situação, tanto pode haver sistemas com uma única solução quanto sistemas com infinitas soluções.

Sistemas Lineares sobre $\mathbb{Z}_p$

Até aqui, todos os sistemas lineares que encontramos envolveram números reais, e as soluções, em concordância, têm sido vetores em algum $\mathbb{R}^n$. Vimos como outros sistemas numéricos podem surgir — notadamente, $\mathbb{Z}_p$. Quando p é um número primo, $\mathbb{Z}_p$ comporta-se como $\mathbb{R}$ sob muitos aspectos; em particular, podemos somar, subtrair, multiplicar e dividir (por números não nulos).[2] Assim, podemos também resolver

[2] $\mathbb{R}$ e $\mathbb{Z}_p$ são exemplos de uma estrutura algébrica chamada *corpo*. O conjunto dos números racionais $\mathbb{Q}$ e o conjunto dos números complexos $\mathbb{C}$ são outros exemplos. Corpos são ensinados com detalhes em cursos de álgebra abstrata.

Material com direitos autorais

sistemas de equações lineares quando as variáveis e os coeficientes pertencem a algum $\mathbb{Z}_p$. Em tais situações, dizemos resolver um sistema *sobre* $\mathbb{Z}_p$.

Por exemplo, a equação linear $x_1 + x_2 + x_3 = 1$, quando vista como uma equação sobre $\mathbb{Z}_2$, tem exatamente quatro soluções:

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \quad \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \quad \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}, \quad \text{e} \quad \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

(A última solução surge porque, em $\mathbb{Z}_2$, $1 + 1 + 1 = 1$.)

Quando olhamos a equação $x_1 + x_2 + x_3 = 1$ sobre $\mathbb{Z}_3$, as soluções $\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$ são

$$\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 2 \\ 2 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 2 \\ 2 \end{bmatrix}, \begin{bmatrix} 2 \\ 0 \\ 2 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix}, \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix} \quad \text{e} \quad \begin{bmatrix} 2 \\ 1 \\ 1 \end{bmatrix}$$

(Verifique.)

Mas nós não precisamos usar métodos de tentativa e erro; os métodos de escalonamento de matrizes completas funcionam tão bem sobre $\mathbb{Z}_p$ como sobre $\mathbb{R}$.

EXEMPLO 10 Resolva os seguintes sistemas de equações lineares sobre $\mathbb{Z}_3$:

$$\begin{aligned} x_1 + 2x_2 + x_3 &= 0 \\ x_1 \qquad + x_3 &= 2 \\ x_2 + 2x_3 &= 1 \end{aligned}$$

SOLUÇÃO: A primeira coisa a notar em exemplos como este é que subtração e divisão não são necessárias; podemos conseguir os mesmos efeitos usando adição e multiplicação. (Isso, contudo, requer que estejamos trabalhando sobre $\mathbb{Z}_p$, onde p é primo; veja o Exercício 58, no final desta seção, e o Exercício 33, na Seção 1.5.)

Vamos agora reduzir a matriz completa do sistema usando cálculos módulo 3.

$$\left[\begin{array}{ccc|c} 1 & 2 & 1 & 0 \\ 1 & 0 & 1 & 2 \\ 0 & 1 & 2 & 1 \end{array}\right] \xrightarrow{L_2 + 2L_1} \left[\begin{array}{ccc|c} 1 & 2 & 1 & 0 \\ 0 & 1 & 0 & 2 \\ 0 & 1 & 2 & 1 \end{array}\right]$$

$$\xrightarrow[L_3 + 2L_2]{L_1 + L_2} \left[\begin{array}{ccc|c} 1 & 0 & 1 & 2 \\ 0 & 1 & 0 & 2 \\ 0 & 0 & 2 & 2 \end{array}\right]$$

$$\xrightarrow[2L_3]{L_1 + L_3} \left[\begin{array}{ccc|c} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 2 \\ 0 & 0 & 1 & 1 \end{array}\right]$$

A solução, portanto, é $x_1 = 1, x_2 = 2, x_3 = 1$. ◆

Material com direitos autorais

EXEMPLO 11 Resolva o seguinte sistema de equações lineares sobre $\mathbb{Z}_2$:

$$\begin{aligned} x_1 + x_2 + x_3 + x_4 &= 1 \\ x_1 + x_2 \qquad\qquad &= 1 \\ x_2 + x_3 \qquad &= 0 \\ x_3 + x_4 &= 0 \\ x_1 \qquad\qquad + x_4 &= 1 \end{aligned}$$

SOLUÇÃO: A redução por linhas é feita da seguinte maneira:

$$\left[\begin{array}{cccc|c} 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 & 1 \end{array}\right] \xrightarrow[L_5 + L_1]{L_2 + L_1} \left[\begin{array}{cccc|c} 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 & 0 \\ 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 \\ 0 & 1 & 1 & 0 & 0 \end{array}\right]$$

$$\xrightarrow[\substack{L_1 + L_2 \\ L_5 + L_2}]{L_2 \leftrightarrow L_3} \left[\begin{array}{cccc|c} 1 & 0 & 0 & 1 & 1 \\ 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right]$$

$$\xrightarrow[L_4 + L_3]{L_2 + L_3} \left[\begin{array}{cccc|c} 1 & 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right]$$

Assim, temos

$$\begin{aligned} x_1 \qquad\qquad + x_4 &= 1 \\ x_2 \qquad + x_4 &= 0 \\ x_3 + x_4 &= 0 \end{aligned}$$

Fixando a variável livre $x_4 = t$, obtemos

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 1 + t \\ t \\ t \\ t \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} + t \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}$$

Como t pode assumir os valores 0 ou 1, existem exatamente duas soluções:

$$\begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} \quad \text{e} \quad \begin{bmatrix} 0 \\ 1 \\ 1 \\ 1 \end{bmatrix}$$

◆

Observação • Sistemas lineares sobre $\mathbb{Z}_p$ nunca podem ter infinitas soluções. (Por que não?) Mais precisamente, quando há mais de uma solução, o número de soluções é finito e é função tanto do número de variáveis livres como de p. (Veja o Exercício 57.)

Material com direitos autorais

◆ EXERCÍCIOS 2.3 ◆

Nos Exercícios de 1 a 8, determine se a matriz dada está na forma escalonada por linhas. Se estiver, indique se ela também está na forma escalonada reduzida.

1. $\begin{bmatrix} 1 & 0 & 1 \\ 0 & 0 & 3 \\ 0 & 1 & 0 \end{bmatrix}$ **2.** $\begin{bmatrix} 7 & 0 & 1 & 0 \\ 0 & 1 & -1 & 4 \\ 0 & 0 & 0 & 0 \end{bmatrix}$

3. $\begin{bmatrix} 0 & 1 & 3 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$ **4.** $\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$

5. $\begin{bmatrix} 1 & 0 & 3 & -4 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 5 & 0 & 1 \end{bmatrix}$ **6.** $\begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}$

7. $\begin{bmatrix} 1 & 2 & 3 \\ 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}$ **8.** $\begin{bmatrix} 2 & 1 & 3 & 5 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 \end{bmatrix}$

Nos Exercícios de 9 a 14, use operações elementares para reduzir a matriz dada às formas (a) escalonada por linhas e (b) escalonada reduzida.

9. $\begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$ **10.** $\begin{bmatrix} 4 & 3 \\ 2 & 1 \end{bmatrix}$

11. $\begin{bmatrix} 3 & 5 \\ 5 & -2 \\ 2 & 4 \end{bmatrix}$ **12.** $\begin{bmatrix} 2 & -4 & -2 & 6 \\ 3 & 1 & 6 & 6 \end{bmatrix}$

13. $\begin{bmatrix} 3 & -2 & -1 \\ 2 & -1 & -1 \\ 4 & -3 & -1 \end{bmatrix}$ **14.** $\begin{bmatrix} -2 & -4 & 7 \\ -3 & -6 & 10 \\ 1 & 2 & -3 \end{bmatrix}$

15. Reverta as operações elementares com as linhas usadas no Exemplo 3 para mostrar que podemos converter

$$\begin{bmatrix} 1 & 2 & -4 & -4 & 5 \\ 0 & -1 & 10 & 9 & -5 \\ 0 & 0 & 1 & 1 & -1 \\ 0 & 0 & 0 & 0 & 24 \end{bmatrix} \text{ em } \begin{bmatrix} 1 & 2 & -4 & -4 & 5 \\ 2 & 4 & 0 & 0 & 2 \\ 2 & 3 & 2 & 1 & 5 \\ -1 & 1 & 3 & 6 & 5 \end{bmatrix}$$

16. Em geral, qual é a operação elementar com linhas que desfaz cada uma das três operações elementares com linhas $L_i \leftrightarrow L_j$, kL_i e $L_i + kL_j$?

Nos Exercícios 17 e 18, mostre que as matrizes dadas são linha-equivalentes e encontre uma seqüência de operações elementares com as linhas que convertem A em B.

17. $A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}, B = \begin{bmatrix} 3 & -1 \\ 1 & 0 \end{bmatrix}$

18. $A = \begin{bmatrix} 2 & 0 & -1 \\ 1 & 1 & 0 \\ -1 & 1 & 1 \end{bmatrix}, B = \begin{bmatrix} 3 & 1 & -1 \\ 3 & 5 & 1 \\ 2 & 2 & 0 \end{bmatrix}$

19. O que está errado com a seguinte demonstração de que toda matriz com pelo menos duas linhas é linha-equivalente a uma matriz com uma linha nula?

Efetue $L_2 + L_1$ e $L_1 + L_2$. Agora as linhas 1 e 2 são idênticas. Efetue agora $L_2 - L_1$ para obter uma linha de zeros na segunda linha.

20. Qual o efeito resultante ao efetuar as seguintes operações elementares sobre linhas em uma matriz (com pelo menos duas linhas)?

$$L_2 + L_1,\ L_1 - L_2,\ L_2 + L_1,\ -L_1$$

21. Qual é o posto de cada uma das matrizes dos Exercícios 1 a 8?

22. Quais são as possíveis formas escalonadas reduzidas de matrizes 3×3?

Nos Exercícios de 23 a 32, resolva o sistema de equações dado usando o método de eliminação de Gauss ou o método de eliminação de Gauss-Jordan.

23. $\begin{aligned} x_1 + 2x_2 - 3x_3 &= 9 \\ 2x_1 - x_2 + x_3 &= 0 \\ 4x_1 - x_2 + x_3 &= 4 \end{aligned}$

24. $\begin{aligned} x - y + z &= 0 \\ -x + 3y + z &= 5 \\ 3x + y + 7z &= 2 \end{aligned}$

25. $\begin{aligned} x_1 - 3x_2 - 2x_3 &= 0 \\ -x_1 + 2x_2 + x_3 &= 0 \\ 2x_1 + 4x_2 + 6x_3 &= 0 \end{aligned}$

26. $\begin{aligned} 2w + 3x - y + 4z &= 0 \\ 3w - x + z &= 1 \\ 3w - 4x + y - z &= 2 \end{aligned}$

27. $\begin{aligned} 2r + s &= 3 \\ 4r + s &= 7 \\ 2r + 5s &= -1 \end{aligned}$

28. $\begin{aligned} -x_1 + 3x_2 - 2x_3 + 4x_4 &= 0 \\ 2x_1 - 6x_2 + x_3 - 2x_4 &= -3 \\ x_1 - 3x_2 + 4x_3 - 8x_4 &= 2 \end{aligned}$

Material com direitos autorais

29. $\frac{1}{2}x_1 + x_2 - x_3 - 6x_4 = 2$

$\frac{1}{6}x_1 + \frac{1}{2}x_2 - 3x_4 + x_5 = -1$

$\frac{1}{3}x_1 - 2x_3 - 4x_5 = 8$

30. $\sqrt{2}x + y + 2z = 1$

$\sqrt{2}y - 3z = -\sqrt{2}$

$-y + \sqrt{2}z = 1$

31. $w + x + 2y + z = 1$

$w - x - y + z = 0$

$x + y = -1$

$w + x + z = 2$

32. $a + b + c + d = 4$

$a + 2b + 3c + 4d = 10$

$a + 3b + 6c + 10d = 20$

$a + 4b + 10c + 20d = 35$

Nos Exercícios de 33 a 36, determine, sem efetuar nenhum cálculo, se um sistema linear com a matriz completa dada tem uma única solução, infinitas soluções ou nenhuma solução. Justifique suas respostas.

33. $\left[\begin{array}{ccc|c} 0 & 0 & 1 & 2 \\ 0 & 1 & 3 & 1 \\ 1 & 0 & 1 & 1 \end{array}\right]$

34. $\left[\begin{array}{cccc|c} 3 & -2 & 0 & 1 & 1 \\ 1 & 2 & -3 & 1 & -1 \\ 2 & 4 & -6 & 2 & 0 \end{array}\right]$

35. $\left[\begin{array}{cccc|c} 1 & 2 & 3 & 4 & 0 \\ 5 & 6 & 7 & 8 & 0 \\ 9 & 10 & 11 & 12 & 0 \end{array}\right]$

36. $\left[\begin{array}{ccccc|c} 1 & 2 & 3 & 4 & 5 & 6 \\ 6 & 5 & 4 & 3 & 2 & 1 \\ 7 & 7 & 7 & 7 & 7 & 7 \end{array}\right]$

37. Mostre que, se $ad - bc \neq 0$, o sistema

$$ax + by = r$$
$$cx + dy = s$$

tem uma única solução.

Nos Exercícios de 38 a 41, para que valor(es) de k, se houver, o sistema terá (a) nenhuma solução, (b) uma única solução e (c) infinitas soluções?

38. $kx + 2y = 3$

$2x - 4y = -6$

39. $x + ky = 1$

$kx + y = 1$

40. $x - 2y + 3z = 2$

$x + y + z = k$

$2x - y + 4z = k^2$

41. $x + y + kz = 1$

$x + ky + z = 1$

$kx + y + z = -2$

42. Dê exemplos de sistemas homogêneos com m equações lineares em n variáveis, com $m = n$ e com $m > n$, que tenham (a) infinitas soluções e (b) uma única solução.

Nos Exercícios 43 e 44, encontre a reta interseção de cada par de planos dados.

43. $3x + 2y + z = -1$ e $2x - y + 4z = 5$

44. $4x + y - z = 0$ e $2x - y + 3z = 4$

45. (a) Dê um exemplo de três planos cuja interseção seja uma reta (Figura 3).

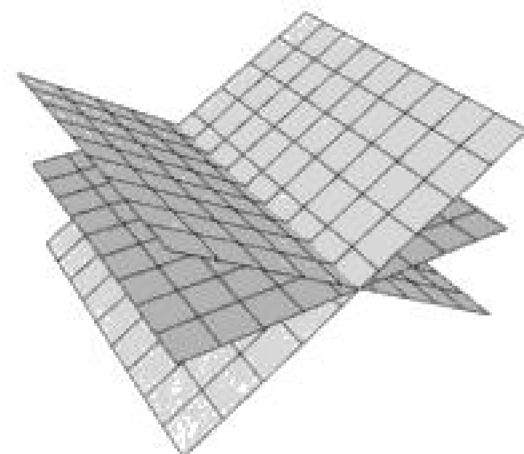


Figura 3

(b) Dê um exemplo de três planos que não tenham nenhum ponto em comum, mas se interceptem dois a dois (Figura 4).

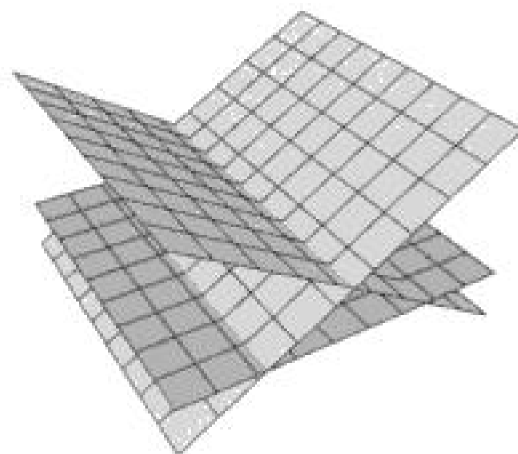


Figura 4

(c) Dê um exemplo de três planos, de modo que exatamente dois deles sejam paralelos (Figura 5).

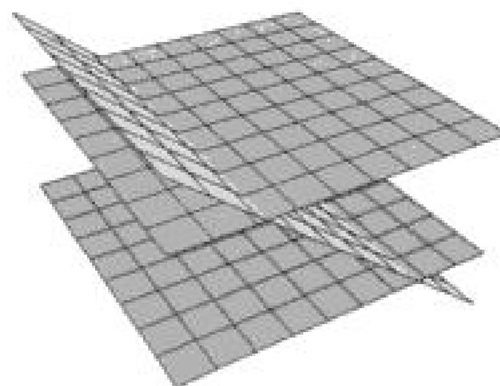


Figura 5

Material com direitos autorais

(d) Dê um exemplo de três planos que se interceptem em um único ponto (Figura 6).

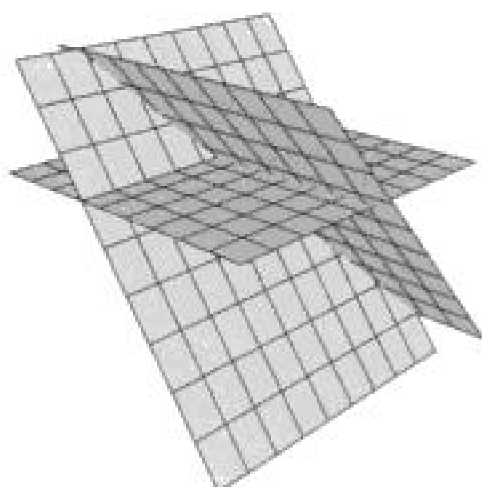


Figura 6

Nos Exercícios 46 e 47, determine se as retas $\mathbf{x} = \mathbf{p} + s\mathbf{u}$ *e* $\mathbf{x} = \mathbf{q} + t\mathbf{v}$ *se interceptam e, nesse caso, encontre o ponto de interseção.*

46. $\mathbf{p} = \begin{bmatrix} -1 \\ 2 \\ 1 \end{bmatrix}, \mathbf{q} = \begin{bmatrix} 2 \\ 2 \\ 0 \end{bmatrix}, \mathbf{u} = \begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix}, \mathbf{v} = \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}$

47. $\mathbf{p} = \begin{bmatrix} 3 \\ 1 \\ 0 \end{bmatrix}, \mathbf{q} = \begin{bmatrix} -1 \\ 1 \\ -1 \end{bmatrix}, \mathbf{u} = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \mathbf{v} = \begin{bmatrix} 2 \\ 3 \\ 1 \end{bmatrix}$

48. Sejam $\mathbf{p} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}, \mathbf{u} = \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix}$, e $\mathbf{v} = \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix}$. Descreva todos os pontos $Q = (a, b, c)$ tais que a reta que passa por Q e tem a direção do vetor $\mathbf{v}$ intercepte a reta de equação $\mathbf{x} = \mathbf{p} + s\mathbf{u}$.

49. Lembre-se de que o produto vetorial dos vetores $\mathbf{u}$ e $\mathbf{v}$ é um vetor $\mathbf{u} \times \mathbf{v}$ ortogonal a ambos os vetores $\mathbf{u}$ e $\mathbf{v}$. (Veja Investigação: O Produto Vetorial, no Capítulo 1.) Se

$$\mathbf{u} = \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix} \quad \text{e} \quad \mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix}$$

mostre que existem infinitos vetores

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

que satisfazem $\mathbf{u} \cdot \mathbf{x} = 0$ e $\mathbf{v} \cdot \mathbf{x} = 0$ simultaneamente e que são múltiplos de $\mathbf{u} \times \mathbf{v} = \begin{bmatrix} u_2v_3 - u_3v_2 \\ u_3v_1 - u_1v_3 \\ u_1v_2 - u_2v_1 \end{bmatrix}$

50. Sejam $\mathbf{p} = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \mathbf{q} = \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix}, \mathbf{u} = \begin{bmatrix} 2 \\ -3 \\ 1 \end{bmatrix}$, e $\mathbf{v} = \begin{bmatrix} 0 \\ 6 \\ -1 \end{bmatrix}$. Mostre que as retas $\mathbf{x} = \mathbf{p} + s\mathbf{u}$ e $\mathbf{x} = \mathbf{q} + t\mathbf{v}$ são reversas. Encontre equações vetoriais de um par de planos paralelos, de modo que cada um deles contenha uma das retas.

Nos Exercícios de 51 a 56, resolva o sistema de equações lineares sobre o corpo $\mathbb{Z}_p$ *indicado.*

51. $\begin{aligned} x + 2y &= 1 \\ x + y &= 2 \end{aligned}$ sobre $\mathbb{Z}_3$

52. $\begin{aligned} x + y &= 1 \\ y + z &= 0 \\ x + z &= 1 \end{aligned}$ sobre $\mathbb{Z}_2$

53. $\begin{aligned} x + y &= 1 \\ y + z &= 0 \\ x + z &= 1 \end{aligned}$ sobre $\mathbb{Z}_3$

54. $\begin{aligned} 3x + 2y &= 1 \\ x + 4y &= 1 \end{aligned}$ sobre $\mathbb{Z}_5$

55. $\begin{aligned} 3x + 2y &= 1 \\ x + 4y &= 1 \end{aligned}$ sobre $\mathbb{Z}_7$

56. $\begin{aligned} x_1 + 4x_4 &= 1 \\ x_1 + 2x_2 + 4x_3 &= 3 \\ 2x_1 + 2x_2 + x_4 &= 1 \\ x_1 + 3x_3 &= 2 \end{aligned}$ sobre $\mathbb{Z}_5$

57. Prove o seguinte corolário do Teorema do Posto: seja A uma matriz $m \times n$ com elementos em $\mathbb{Z}_p$. Qualquer sistema de equações lineares possível, com matriz dos coeficientes igual a A, tem exatamente $p^{n-\text{posto}(A)}$ soluções sobre $\mathbb{Z}_p$.

58. Quando p não é primo, um cuidado extra é necessário ao resolvermos um sistema linear (ou, na verdade, qualquer equação) sobre $\mathbb{Z}_p$. Usando o método de eliminação de Gauss, resolva o seguinte sistema sobre $\mathbb{Z}_6$. Qual complicação surge?

$$\begin{aligned} 2x + 3y &= 4 \\ 4x + 3y &= 2 \end{aligned}$$

Material com direitos autorais

Investigação

Pivoteamento Parcial

Em Investigação: Mentiras que meu Computador me Contou, logo após a Seção 2.2, vimos que sistemas lineares mal-condicionados podem causar problemas quando erros de arredondamento ocorrem. Nesta investigação, você descobrirá um outro tipo de problema no qual sistemas lineares são sensíveis a erros de arredondamento, e verá que mudanças muito pequenas nos coeficientes podem levar a enormes imprecisões na solução. Felizmente, há uma coisa que se pode fazer para minimizar ou mesmo eliminar esse problema (ao contrário do problema com sistemas mal-condicionados).

1. (a) Resolva a equação linear $0{,}00021x = 1$ em x.
   (b) Suponha que sua calculadora possa armazenar apenas quatro algarismos significativos. A equação será arredondada para $0{,}0002x = 1$. Resolva essa equação.

A diferença entre as respostas das partes (a) e (b) pode ser pensada como os efeitos de um erro de 0,00001 na solução da equação dada.

2. Agora estenda essa idéia a um sistema de equações lineares.
   (a) Usando o método de eliminação de Gauss, resolva o sistema linear

$$\begin{aligned} 0{,}400x + 99{,}6y &= 100 \\ 75{,}3x - 45{,}3y &= 30{,}0 \end{aligned}$$

   usando três algarismos significativos. Comece escolhendo 0,400 como pivô e faça cada conta usando três algarismos significativos. Você deverá obter a "solução" $x = -1{,}00$, $y = 1{,}01$. Verifique que a resposta verdadeira é $x = 1{,}00$, $y = 1{,}00$. Esse é um erro enorme — 200% no valor de x! Você pode descobrir o que o causou?
   (b) Resolva o sistema da parte (a) novamente, desta vez trocando de posição as duas equações (ou, equivalentemente, as duas linhas de sua matriz completa) e escolhendo 75,3 como pivô. De novo, use três algarismos significativos. Qual é a solução desta vez?

A moral desta história é que, quando usar o método de eliminação de Gauss ou de Gauss-Jordan para obter uma solução numérica para um sistema de equações lineares, você deve escolher os pivôs com cuidado. Especificamente, em cada passo do pivoteamento, escolha, dentre todos os possíveis pivôs em uma coluna, o elemento com o maior valor absoluto. Use a troca de linhas para levar esse elemento para a posição correspondente e use-o para criar zeros na coluna, onde necessário. Essa estratégia é conhecida como ***pivoteamento parcial***.

3. Resolva os seguintes sistemas pelo método de eliminação de Gauss, primeiramente sem e depois com o pivoteamento parcial. Efetue cada cálculo usando três algarismos significativos. (As soluções exatas são dadas.)

(a)
$$\begin{aligned} 0{,}001x + 0{,}995y &= 1{,}00 \\ -10{,}2x + 1{,}00y &= -50{,}0 \end{aligned}$$

Solução exata: $\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 5{,}00 \\ 1{,}00 \end{bmatrix}$

(b)
$$\begin{aligned} 10x - 7y \phantom{{}+6z} &= 7 \\ -3x + 2{,}09y + 6z &= 3{,}91 \\ 5x - y + 5z &= 6 \end{aligned}$$

Solução exata: $\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0{,}00 \\ -1{,}00 \\ 1{,}00 \end{bmatrix}$

Material com direitos autorais





que pode ser simplificada para

$$\lceil 1 \rceil \; \lceil -1 \rceil$$

(c) Mostre que o número de operações necessárias para completar a fase de substituição de trás para a frente é

$$1 + 2 + \cdots + (n - 1)$$

(d) Usando fórmulas de somatórios para as somas nas partes (b) e (c) (veja os Exercícios 45 e 46, na Seção 2.5, e o Apêndice B), mostre que o número total de operações, $T(n)$, realizadas no método de eliminação de Gauss é

$$T(n) = \tfrac{1}{3}n^3 + n^2 - \tfrac{1}{3}n$$

Como toda função polinomial é aproximada, para valores grandes da variável, por seu termo dominante, vemos que $T(n) \approx \frac{1}{3}n^3$ para valores grandes de n.

4. Mostre que o método de eliminação de Gauss-Jordan terá um total de $T(n) \approx \frac{1}{2}n^3$ operações se criarmos zeros acima e abaixo do 1 líder conforme efetuamos os passos. (Isso mostra que, para sistemas grandes de equações lineares, o método de eliminação de Gauss é mais rápido que a versão de Gauss-Jordan.)

Material com direitos autorais

2.4 Conjuntos Geradores e Dependência Linear

O segundo dos três caminhos em nossas "três vias" do estudo de sistemas de equações lineares é voltado para as combinações lineares de vetores. Vimos que podemos encarar a resolução de um sistema de equações lineares como uma questão sobre a possibilidade de expressar um vetor como combinação linear de outros vetores dados. Vamos explorar essa idéia em mais detalhes nesta seção. Ela conduz a alguns conceitos muito importantes, que serão encontrados repetidamente em capítulos posteriores.

Conjuntos Geradores

Podemos agora responder facilmente à questão levantada na Seção 1.2: quando um dado vetor é combinação linear de outros vetores dados?

EXEMPLO 1 (a) O vetor $\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$ é uma combinação linear dos vetores $\begin{bmatrix} 1 \\ 0 \\ 3 \end{bmatrix}$ e $\begin{bmatrix} -1 \\ 1 \\ -3 \end{bmatrix}$?

(b) O vetor $\begin{bmatrix} 2 \\ 3 \\ 4 \end{bmatrix}$ é uma combinação linear dos vetores $\begin{bmatrix} 1 \\ 0 \\ 3 \end{bmatrix}$ e $\begin{bmatrix} -1 \\ 1 \\ -3 \end{bmatrix}$?

SOLUÇÃO: (a) Queremos encontrar escalares x e y tais que

$$x\begin{bmatrix} 1 \\ 0 \\ 3 \end{bmatrix} + y\begin{bmatrix} -1 \\ 1 \\ -3 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$$

Desenvolvendo, obtemos o sistema

$$\begin{aligned} x - y &= 1 \\ y &= 2 \\ 3x - 3y &= 3 \end{aligned}$$

cuja matriz completa é

$$\left[\begin{array}{rr|r} 1 & -1 & 1 \\ 0 & 1 & 2 \\ 3 & -3 & 3 \end{array}\right]$$

(Observe que as colunas da matriz completa são justamente os vetores dados; repare na ordem dos vetores — em particular, qual vetor é o vetor constante.)

A forma escalonada reduzida dessa matriz é

$$\left[\begin{array}{rr|r} 1 & 0 & 3 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \end{array}\right]$$

Material com direitos autorais

(Verifique esse fato.) Assim, a solução é $x = 3$, $y = 2$, e a combinação linear correspondente é

$$3\begin{bmatrix}1\\0\\3\end{bmatrix} + 2\begin{bmatrix}-1\\1\\-3\end{bmatrix} = \begin{bmatrix}1\\2\\3\end{bmatrix}$$

(b) Utilizando nossa observação feita no item (a), obtemos um sistema linear cuja matriz completa é

$$\left[\begin{array}{cc|c}1 & -1 & 2\\0 & 1 & 3\\3 & -3 & 4\end{array}\right]$$

que se reduz a

$$\left[\begin{array}{cc|c}1 & 0 & 5\\0 & 1 & 3\\0 & 0 & -2\end{array}\right]$$

revelando que o sistema não tem solução. Assim, neste caso, $\begin{bmatrix}2\\3\\4\end{bmatrix}$ não é uma combinação linear de $\begin{bmatrix}1\\0\\3\end{bmatrix}$ e $\begin{bmatrix}-1\\1\\-3\end{bmatrix}$. 

O problema de verificar se um conjunto é gerador está intimamente relacionado com o problema da existência de solução de sistemas lineares. Veja novamente o Exemplo 1. Nele, vimos que um sistema com matriz completa $[A \mid \mathbf{b}]$ tem solução exatamente quando $\mathbf{b}$ é uma combinação linear das colunas de A. Esse é um fato geral, resumido no próximo teorema.

◆ TEOREMA 1

Um sistema de equações lineares com matriz completa $[A \mid \mathbf{b}]$ é possível se, e somente se, $\mathbf{b}$ for uma combinação linear das colunas de A.

Vamos revisitar o Exemplo 4 da Seção 2.2, interpretando-o à luz do Teorema 1.

(a) O sistema

$$x - y = 1$$
$$x + y = 3$$

tem uma única solução $x = 2$, $y = 1$. Assim,

$$2\begin{bmatrix}1\\1\end{bmatrix} + \begin{bmatrix}-1\\1\end{bmatrix} = \begin{bmatrix}1\\3\end{bmatrix}$$

Veja a Figura 1(a).

Material com direitos autorais

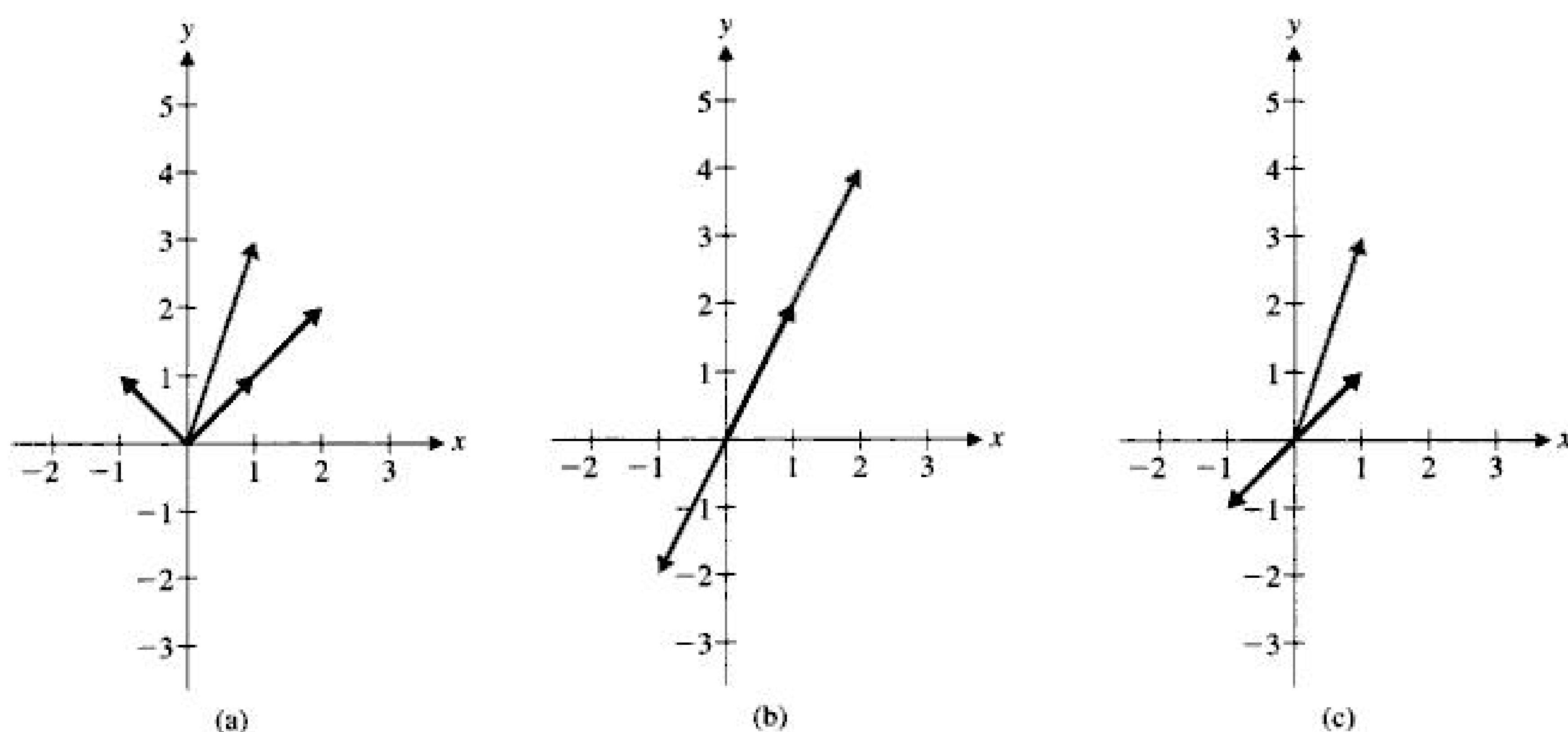


Figura 1

(b) O sistema

$$x - y = 2$$
$$2x - 2y = 4$$

tem infinitas soluções da forma $x = 2 + t$, $y = t$. Isso significa que

$$(2 + t)\begin{bmatrix} 1 \\ 2 \end{bmatrix} + t\begin{bmatrix} -1 \\ -2 \end{bmatrix} = \begin{bmatrix} 2 \\ 4 \end{bmatrix}$$

para todos os valores de t. Geometricamente, os vetores $\begin{bmatrix} 1 \\ 2 \end{bmatrix}$, $\begin{bmatrix} -1 \\ -2 \end{bmatrix}$ e $\begin{bmatrix} 2 \\ 4 \end{bmatrix}$ são todos paralelos, por isso todos estão alinhados ao longo da mesma reta que passa pela origem [veja a Figura 1(b)].

(c) O sistema

$$x - y = 1$$
$$x - y = 3$$

não tem soluções, e, portanto, não existem valores x e y que satisfazem

$$x\begin{bmatrix} 1 \\ 1 \end{bmatrix} + y\begin{bmatrix} -1 \\ -1 \end{bmatrix} = \begin{bmatrix} 1 \\ 3 \end{bmatrix}$$

Nesse caso, $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$ e $\begin{bmatrix} -1 \\ -1 \end{bmatrix}$ são paralelos, mas $\begin{bmatrix} 1 \\ 3 \end{bmatrix}$ não está alinhado ao longo da mesma reta que passa pela origem e sobre a qual se alinham os outros vetores [veja a Figura 1(c)].

Estaremos freqüentemente interessados na coleção de *todas* as combinações lineares de um dado conjunto de vetores.

Material com direitos autorais

Definição Se $S = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\}$ é um conjunto de vetores em $\mathbb{R}^n$, o conjunto de todas as combinações lineares de $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k$ é chamado ***conjunto gerado por*** $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k$ e é denotado por $\text{ger}(\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k)$, ou $\text{ger}(S)$. Se $\text{ger}(S) = \mathbb{R}^n$, S é chamado de ***conjunto gerador*** de $\mathbb{R}^n$.

EXEMPLO 2 Mostre que $\mathbb{R}^2 = \text{ger}\left(\begin{bmatrix} 2 \\ -1 \end{bmatrix}, \begin{bmatrix} 1 \\ 3 \end{bmatrix}\right)$.

SOLUÇÃO: Precisamos mostrar que um vetor arbitrário $\begin{bmatrix} a \\ b \end{bmatrix}$ pode ser escrito como uma combinação linear de $\begin{bmatrix} 2 \\ -1 \end{bmatrix}$ e $\begin{bmatrix} 1 \\ 3 \end{bmatrix}$; precisamos, portanto, mostrar que a equação $x\begin{bmatrix} 2 \\ -1 \end{bmatrix} + y\begin{bmatrix} 1 \\ 3 \end{bmatrix} = \begin{bmatrix} a \\ b \end{bmatrix}$ pode sempre ser resolvida para x e y (em termos de a e b), independentemente dos valores de a e b.

A matriz completa é $\left[\begin{array}{cc|c} 2 & 1 & a \\ -1 & 3 & b \end{array}\right]$, e a redução por linhas produz

$$\left[\begin{array}{cc|c} 2 & 1 & a \\ -1 & 3 & b \end{array}\right] \xrightarrow{L_1 \leftrightarrow L_2} \left[\begin{array}{cc|c} -1 & 3 & b \\ 2 & 1 & a \end{array}\right] \xrightarrow{L_2 + 2L_1} \left[\begin{array}{cc|c} -1 & 3 & b \\ 0 & 7 & a + 2b \end{array}\right]$$

de modo que fica claro que o sistema tem uma (única) solução. (Por quê?) Se continuarmos, obteremos

$$\xrightarrow{\frac{1}{7}L_2} \left[\begin{array}{cc|c} -1 & 3 & b \\ 0 & 1 & (a + 2b)/7 \end{array}\right] \xrightarrow{L_1 - 3L_2} \left[\begin{array}{cc|c} -1 & 0 & (b - 3a)/7 \\ 0 & 1 & (a + 2b)/7 \end{array}\right]$$

de onde veremos que $x = (3a - b)/7$ e $y = (a + 2b)/7$. Assim, para qualquer escolha de a e b, teremos

$$\left(\frac{3a - b}{7}\right)\begin{bmatrix} 2 \\ -1 \end{bmatrix} + \left(\frac{a + 2b}{7}\right)\begin{bmatrix} 1 \\ 3 \end{bmatrix} = \begin{bmatrix} a \\ b \end{bmatrix}$$

(Verifique isso.) ◆

Observação: É verdade também que $\mathbb{R}^2 = \text{ger}\left(\begin{bmatrix} 2 \\ -1 \end{bmatrix}, \begin{bmatrix} 1 \\ 3 \end{bmatrix}, \begin{bmatrix} 5 \\ 7 \end{bmatrix}\right)$. Se, dado $\begin{bmatrix} a \\ b \end{bmatrix}$, pudermos encontrar x e y tais que $x\begin{bmatrix} 2 \\ -1 \end{bmatrix} + y\begin{bmatrix} 1 \\ 3 \end{bmatrix} = \begin{bmatrix} a \\ b \end{bmatrix}$, então nós também teremos $x\begin{bmatrix} 2 \\ -1 \end{bmatrix} + y\begin{bmatrix} 1 \\ 3 \end{bmatrix} + 0\begin{bmatrix} 5 \\ 7 \end{bmatrix} = \begin{bmatrix} a \\ b \end{bmatrix}$. De fato, qualquer conjunto de vetores que *contenha* um conjunto gerador de $\mathbb{R}^2$ será também um conjunto gerador de $\mathbb{R}^2$ (veja o Exercício 20).

O próximo exemplo é um caso importante (e fácil) de um conjunto gerador. Encontraremos versões dele muitas vezes.

EXEMPLO 3 Sejam $\mathbf{e}_1$, $\mathbf{e}_2$ e $\mathbf{e}_3$ os vetores unitários canônicos em $\mathbb{R}^3$. Para qualquer vetor $\begin{bmatrix} x \\ y \\ z \end{bmatrix}$, temos

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = x\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + y\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} + z\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = x\mathbf{e}_1 + y\mathbf{e}_2 + z\mathbf{e}_3$$

Portanto, $\mathbb{R}^3 = \text{ger}(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3)$.

Você não deverá ter dificuldade para verificar que, em geral, $\mathbb{R}^n = \text{ger}(\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n)$. ◆

Material com direitos autorais





e podemos multiplicar ambos os lados da equação por $1/c_1$ para obter $\mathbf{v}_1$ como uma combinação linear dos outros vetores:

$$\mathbf{v}_1 = -\left(\frac{c_2}{c_1}\right)\mathbf{v}_2 - \cdots - \left(\frac{c_m}{c_1}\right)\mathbf{v}_m$$ ◆

Nota: Pode parecer que estamos trapaceando um pouco nessa demonstração. Afinal de contas, não sabemos ao certo se $\mathbf{v}_1$ é uma combinação linear dos outros vetores, nem se c_1 é não nulo. Entretanto, o argumento é análogo para algum outro vetor $\mathbf{v}_i$ ou para um escalar c_j. Podemos, alternativamente, apenas renomear os índices de modo que eles funcionem como na demonstração. Em uma situação como essa, um matemático pode começar dizendo "Sem perda de generalidade, podemos assumir que $\mathbf{v}_1$ é uma combinação linear dos outros vetores", e então continuar como acabamos de ver.

EXEMPLO 5 Qualquer conjunto de vetores que contenha o vetor nulo é linearmente dependente. Se $\mathbf{0}, \mathbf{v}_2, \ldots, \mathbf{v}_m$ estão em $\mathbb{R}^n$, podemos encontrar uma combinação linear não trivial da forma $c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \ldots + c_m\mathbf{v}_m = \mathbf{0}$ colocando $c_1 = 1$ e $c_2 = c_3 = \ldots = c_m = 0$. ◆

EXEMPLO 6 Determine se os seguintes conjuntos de vetores são linearmente independentes:

(a) $\begin{bmatrix}1\\4\end{bmatrix}$ e $\begin{bmatrix}-1\\2\end{bmatrix}$ (b) $\begin{bmatrix}1\\1\\0\end{bmatrix}, \begin{bmatrix}0\\1\\1\end{bmatrix}$ e $\begin{bmatrix}1\\0\\1\end{bmatrix}$

(c) $\begin{bmatrix}1\\-1\\0\end{bmatrix}, \begin{bmatrix}0\\1\\-1\end{bmatrix}$ e $\begin{bmatrix}-1\\0\\1\end{bmatrix}$ (d) $\begin{bmatrix}1\\2\\0\end{bmatrix}, \begin{bmatrix}1\\1\\-1\end{bmatrix}$ e $\begin{bmatrix}1\\4\\2\end{bmatrix}$

SOLUÇÃO: Ao responder a qualquer questão desse tipo, uma boa idéia é ver se você pode determinar por inspeção se um dos vetores é uma combinação linear dos outros. Com um pouco de observação, é possível economizar muitas contas!

(a) O único modo de dois vetores serem linearmente dependentes é quando um é múltiplo do outro. (Por quê?) Os vetores dados obviamente não são múltiplos, logo, são linearmente independentes.

(b) Não há uma relação de dependência óbvia neste caso, por isso, tentamos encontrar escalares c_1, c_2, c_3 tais que

$$c_1\begin{bmatrix}1\\1\\0\end{bmatrix} + c_2\begin{bmatrix}0\\1\\1\end{bmatrix} + c_3\begin{bmatrix}1\\0\\1\end{bmatrix} = \begin{bmatrix}0\\0\\0\end{bmatrix}$$

O sistema linear correspondente é

$$\begin{aligned} c_1 + \ c_3 &= 0 \\ c_1 + c_2 &= 0 \\ c_2 + c_3 &= 0 \end{aligned}$$

e a matriz completa é

$$\left[\begin{array}{ccc|c} 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \end{array}\right]$$

Novamente, fazemos a observação fundamental de que as colunas dos coeficientes da matriz são exatamente os vetores em questão!

Material com direitos autorais

A forma escalonada reduzida é

$$\left[\begin{array}{ccc|c} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{array}\right]$$

(verifique), logo, $c_1 = 0$, $c_2 = 0$, $c_3 = 0$. Portanto, os vetores dados são linearmente independentes.

(c) Uma pequena reflexão revela que

$$\begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix} + \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

e, portanto, os três vetores são linearmente dependentes. [Monte um sistema linear como na parte (b) para verificar isso algebricamente.]

(d) Novamente, observamos que não há dependência óbvia. Logo, passamos imediatamente à redução a um sistema linear homogêneo cuja matriz completa tem como colunas os vetores dados:

$$\left[\begin{array}{ccc|c} 1 & 1 & 1 & 0 \\ 2 & 1 & 4 & 0 \\ 0 & -1 & 2 & 0 \end{array}\right] \xrightarrow{L_2 - 2L_1} \left[\begin{array}{ccc|c} 1 & 1 & 1 & 0 \\ 0 & -1 & 2 & 0 \\ 0 & -1 & 2 & 0 \end{array}\right] \xrightarrow[\substack{L_3 - L_2 \\ -L_2}]{L_1 + L_2} \left[\begin{array}{ccc|c} 1 & 0 & 3 & 0 \\ 0 & 1 & -2 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right]$$

Se chamarmos os escalares de c_1, c_2 e c_3, teremos

$$\begin{aligned} c_1 + \qquad 3c_3 &= 0 \\ c_2 - 2c_3 &= 0 \end{aligned}$$

de onde vemos que o sistema tem infinitas soluções. Em particular, deverá existir uma solução não nula e, portanto, os vetores dados são linearmente dependentes.
Se continuarmos, poderemos descrever essas soluções exatamente: $c_1 = -3c_3$ e $c_2 = 2c_3$. Logo, para qualquer valor não nulo de c_3, temos a relação de dependência linear

$$-3c_3\begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix} + 2c_3\begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix} + c_3\begin{bmatrix} 1 \\ 4 \\ 2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

(Mais uma vez, verifique que isso está correto.) ◆

Resumimos esse procedimento de verificação de dependência linear no teorema a seguir.

◆ TEOREMA 3

Sejam $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_m$ vetores (coluna) em $\mathbb{R}^n$ e A a matriz $m \times n$ $[\mathbf{v}_1\ \mathbf{v}_2\ \ldots\ \mathbf{v}_m]$ que tem esses vetores como suas colunas. Então, $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_m$ serão linearmente dependentes se, e somente se, o sistema linear homogêneo cuja matriz completa for $[A \mid \mathbf{0}]$ tiver uma solução não trivial.

Material com direitos autorais

DEMONSTRAÇÃO: Os vetores $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_m$ serão linearmente dependentes se, e somente se, existirem escalares $c_1, c_2, \ldots c_m$, não todos nulos, tais que $c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \ldots + c_m\mathbf{v}_m = \mathbf{0}$. Pelo Teorema 1, isso é equivalente a dizer que o vetor não nulo $\begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_m \end{bmatrix}$ é uma solução do sistema cuja matriz completa é $[\mathbf{v}_1\ \mathbf{v}_2 \cdots \mathbf{v}_m \mid \mathbf{0}]$. ◆

EXEMPLO 7 Os vetores unitários canônicos $\mathbf{e}_1$, $\mathbf{e}_2$, e $\mathbf{e}_3$ são linearmente independentes em $\mathbb{R}^3$, pois o sistema cuja matriz completa é $[\mathbf{e}_1\ \mathbf{e}_2\ \mathbf{e}_3 \mid \mathbf{0}]$ já está na forma escalonada reduzida, portanto,

$$\left[\begin{array}{ccc|c} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{array}\right]$$

claramente admite apenas a solução trivial. Em geral, vemos que $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n$ será linearmente independente em $\mathbb{R}^n$. ◆

A realização de operações elementares com linhas em uma matriz constrói combinações lineares das linhas. Podemos usar esse fato para trazer mais um modo de testar a dependência linear de vetores.

EXEMPLO 8 Considere os três vetores do Exemplo 6(d) como vetores-*linha*:

$$[1, 2, 0], [1, 1, -1] \text{ e } [1, 4, 2]$$

Construímos uma matriz tendo esses vetores como suas linhas e passamos ao escalonamento. Cada vez que uma linha muda, denotamos a nova linha adicionando um apóstrofo:

$$\begin{bmatrix} 1 & 2 & 0 \\ 1 & 1 & -1 \\ 1 & 4 & 2 \end{bmatrix} \xrightarrow[L_3' = L_3 - L_1]{L_2' = L_2 - L_1} \begin{bmatrix} 1 & 2 & 0 \\ 0 & -1 & -1 \\ 0 & 2 & 2 \end{bmatrix} \xrightarrow{L_3'' = L_3' + 2L_2'} \begin{bmatrix} 1 & 2 & 0 \\ 0 & -1 & -1 \\ 0 & 0 & 0 \end{bmatrix}$$

Disso, vemos que

$$\mathbf{0} = L_3'' = L_3' + 2L_2' = (L_3 - L_1) + 2(L_2 - L_1) = -3L_1 + 2L_2 + L_3$$

ou, em termos dos vetores originais,

$$-3[1, 2, 0] + 2[1, 1, -1] + [1, 4, 2] = [0, 0, 0]$$

[Note que essa abordagem corresponde a tomar $c_3 = 1$ na solução do Exemplo 6(d).] ◆

As linhas de uma matriz, portanto, serão linearmente dependentes se operações elementares com linhas puderem ser usadas para criar uma linha nula. Resumimos essa descoberta da seguinte maneira:

◆ **TEOREMA 4**

Sejam $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_m$ vetores (linha) em $\mathbb{R}^n$ e A a matriz $m \times n$ $\begin{bmatrix} \mathbf{v}_1 \\ \mathbf{v}_2 \\ \vdots \\ \mathbf{v}_m \end{bmatrix}$ que tem esses vetores como suas linhas. Então, $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_m$ serão linearmente dependentes se, e somente se, $\text{posto}(A) < m$.

Material com direitos autorais











Análise de Redes

Redes aparecem em várias situações práticas: redes de transporte, redes de comunicação e redes econômicas, para mencionar algumas. São particularmente interessantes os possíveis *fluxos* através de redes. Por exemplo, veículos fluem através de redes de estradas, informação flui através de uma rede de dados, bens e serviços fluem através de uma rede econômica.

Para nós, uma ***rede*** consiste em um número finito de ***nós*** (também chamados ***junções*** ou ***vértices***) conectados por uma série de segmentos dirigidos, conhecidos como ***ramos*** ou ***arcos***. Cada ramo é rotulado com um ***fluxo*** que representa a quantidade de alguma mercadoria que pode fluir ao longo ou através daquele ramo na direção indicada. (Pense em carros viajando ao longo de uma rede de ruas de mão única.) A regra fundamental que governa o fluxo através da rede é a ***conservação de fluxo***:

> Em cada nó, o fluxo de entrada é igual ao fluxo de saída.

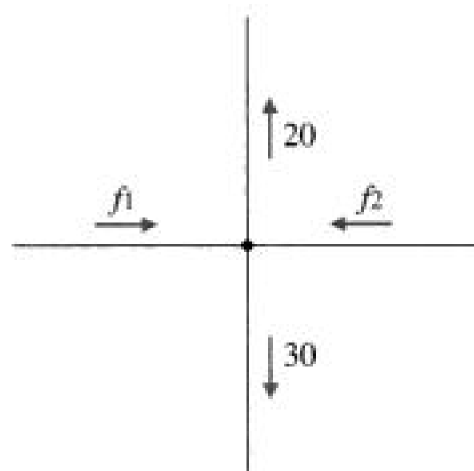


Figura 1 Fluxo em um nó: $f_1 + f_2 = 50$

A Figura 1 mostra uma parte de uma rede, com dois ramos entrando em um nó e dois saindo. A regra de conservação do fluxo implica que o fluxo total de entrada, $f_1 + f_2$ unidades, deve coincidir com o fluxo total de saída, 20 + 30 unidades. Assim, temos a equação linear $f_1 + f_2 = 50$ correspondente a esse nó.

Podemos analisar o fluxo através de uma rede inteira construindo tais equações e resolvendo o sistema de equações lineares resultante.

EXEMPLO 4 Descreva os possíveis fluxos através da rede de encanamento de água mostrada na Figura 2, onde o fluxo é medido em litros por minuto.

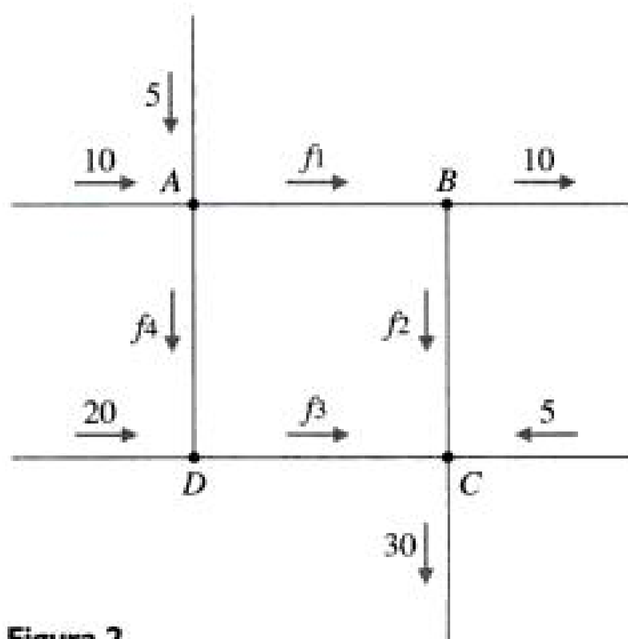


Figura 2

Material com direitos autorais



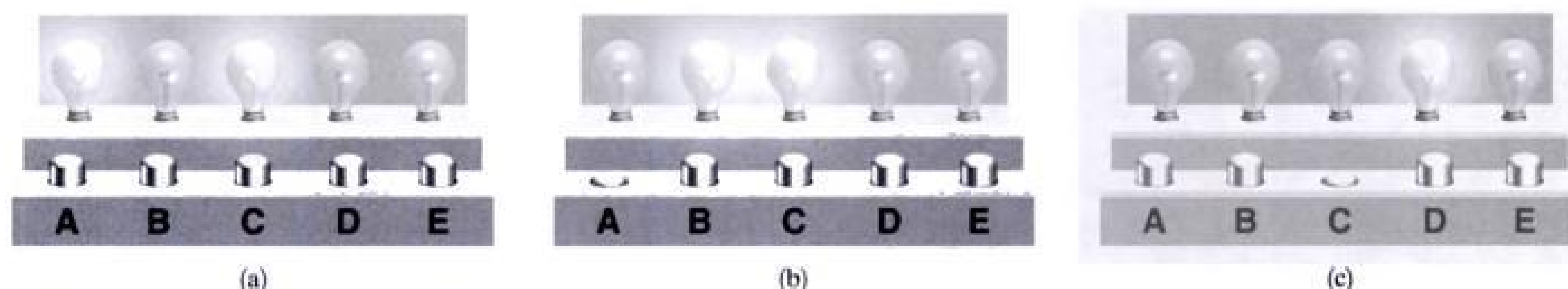






e um interruptor pode mudar o estado da lâmpada de acesa para apagada e vice-versa. Sistemas digitais que surgem em ciência da computação muitas vezes são desse tipo. Muitos jogos de computador retratam quebra-cabeças nos quais um certo esquema deve ser manipulado por vários interruptores a fim de produzir o efeito desejado. A natureza finita de tais situações é perfeitamente adequada a análises que fazem uso de aritmética modular, e freqüentemente sistemas lineares sobre $\mathbb{Z}_p$ entram em cena. Problemas que envolvem esse tipo de situação são chamados ***jogos lineares finitos***.

EXEMPLO 7 Uma fileira de cinco lâmpadas é controlada por cinco interruptores. Cada interruptor muda o estado (ligado ou desligado) da lâmpada diretamente sobre ele e os estados das lâmpadas imediatamente adjacentes à esquerda e à direita.



(a) (b) (c)

Figura 6

Por exemplo, se a primeira e a terceira lâmpadas estão acesas, como na Figura 6(a), ao ser pressionado o interruptor A, muda-se o estado do sistema para o mostrado na Figura 6(b). Se depois pressionarmos o interruptor C, o resultado será o mostrado na Figura 6(c).

Suponha inicialmente que todas as luzes estejam apagadas. Podemos pressionar os interruptores em alguma ordem de modo que a primeira, a terceira e a quinta lâmpadas fiquem acesas? Podemos pressionar os interruptores em alguma ordem de modo que só a primeira lâmpada fique acesa?

SOLUÇÃO: A natureza liga/desliga deste problema sugere que a notação binária será útil e que deveremos trabalhar em $\mathbb{Z}_2$. Assim, representamos os estados das cinco lâmpadas por um vetor em $\mathbb{Z}_2^5$, de modo que 0 representa desligado e 1, ligado. Por exemplo, o vetor

$$\begin{bmatrix} 0 \\ 1 \\ 1 \\ 0 \\ 0 \end{bmatrix}$$

corresponde à Figura 6(b).

Podemos também usar vetores em $\mathbb{Z}_2^5$ para representar a ação de cada interruptor. Se um interruptor muda o estado de uma lâmpada, a componente correspondente é um 1; caso contrário, ela é 0. Com essa convenção, as ações dos cinco interruptores são dadas por

$$\mathbf{a} = \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \quad \mathbf{c} = \begin{bmatrix} 0 \\ 1 \\ 1 \\ 1 \\ 0 \end{bmatrix}, \quad \mathbf{d} = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 1 \\ 1 \end{bmatrix}, \quad \mathbf{e} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 1 \end{bmatrix}$$

A situação ilustrada na Figura 6(a) corresponde ao estado inicial

Material com direitos autorais

$$\mathbf{s} = \begin{bmatrix} 1 \\ 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}$$

seguida por

$$\mathbf{a} = \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

A soma (em $\mathbb{Z}_2^5$) é

$$\mathbf{s} + \mathbf{a} = \begin{bmatrix} 0 \\ 1 \\ 1 \\ 0 \\ 0 \end{bmatrix}$$

Observe que esse resultado corresponde à situação encontrada na Figura 6(b).

Começando com qualquer configuração $\mathbf{s}$, suponha que pressionemos os interruptores na ordem A, C, D, A, C, B. Isso corresponde à soma vetorial $\mathbf{s} + \mathbf{a} + \mathbf{c} + \mathbf{d} + \mathbf{a} + \mathbf{c} + \mathbf{b}$. Como em $\mathbb{Z}_2^5$ a adição é comutativa, temos

$$\mathbf{s} + \mathbf{a} + \mathbf{c} + \mathbf{d} + \mathbf{a} + \mathbf{c} + \mathbf{b} = \mathbf{s} + 2\mathbf{a} + \mathbf{b} + 2\mathbf{c} + \mathbf{d} = \mathbf{s} + \mathbf{b} + \mathbf{d}$$

já que $2 = 0$ em $\mathbb{Z}_2$. Assim, poderíamos chegar ao mesmo resultado pressionando apenas B e D — não importando a ordem. (Verifique que essa afirmação está correta.) Logo, neste exemplo, não precisamos pressionar nenhum interruptor mais de uma vez.

Para decidir se é possível chegar a uma configuração-alvo $\mathbf{t}$, começando de uma configuração inicial $\mathbf{s}$, precisamos determinar se existem escalares $x_1, \ldots, x_5$ em $\mathbb{Z}_2$ tais que

$$\mathbf{s} + x_1\mathbf{a} + x_2\mathbf{b} + \ldots + x_5\mathbf{e} = \mathbf{t}$$

Em outras palavras, precisamos resolver (se possível) o sistema linear sobre $\mathbb{Z}_2$ que corresponde à equação vetorial

$$x_1\mathbf{a} + x_2\mathbf{b} + \ldots + x_5\mathbf{e} = \mathbf{t} - \mathbf{s} = \mathbf{t} + \mathbf{s}$$

Neste problema, $\mathbf{s} = \mathbf{0}$, e nossa primeira configuração-alvo é

$$\mathbf{t} = \begin{bmatrix} 1 \\ 0 \\ 1 \\ 0 \\ 1 \end{bmatrix}$$

A matriz completa desse sistema tem os vetores dados como colunas:

$$\left[\begin{array}{ccccc|c} 1 & 1 & 0 & 0 & 0 & 1 \\ 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 \end{array}\right]$$

Material com direitos autorais

Reduzindo-a sobre $\mathbb{Z}_2$, obtemos

$$\left[\begin{array}{ccccc|c} 1 & 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right]$$

Vemos que x_5 é uma variável livre. Portanto, há exatamente duas soluções (correspondentes a $x_5 = 0$ e $x_5 = 1$). Resolvendo o sistema para as outras variáveis em termos de x_5, obtemos

$$\begin{aligned} x_1 &= x_5 \\ x_2 &= 1 + x_5 \\ x_3 &= 1 \\ x_4 &= 1 + x_5 \end{aligned}$$

Logo, quando $x_5 = 0$ e $x_5 = 1$, temos respectivamente as soluções

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ 1 \\ 1 \\ 0 \end{bmatrix} \quad \text{e} \quad \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 1 \\ 0 \\ 1 \end{bmatrix}$$

(Verifique que essas duas soluções funcionam.)
Analogamente, no segundo caso, temos

$$\mathbf{t} = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

A matriz completa se reduz da seguinte maneira:

$$\left[\begin{array}{ccccc|c} 1 & 1 & 0 & 0 & 0 & 1 \\ 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 \end{array}\right] \longrightarrow \left[\begin{array}{ccccc|c} 1 & 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{array}\right]$$

mostrando que não há solução neste caso; é impossível começar com todas as lâmpadas apagadas e acender apenas a primeira. ◆

O Exemplo 7 mostra o poder da álgebra linear. Embora pudéssemos ter descoberto por tentativa e erro que não havia solução, testar todas as possíveis maneiras de ligar os interruptores teria sido extremamente tedioso. Além disso, poderíamos ter deixado de notar o fato de que nenhum dos interruptores precisa ser acionado mais do que uma vez.

EXEMPLO 8 Considere uma fileira com apenas três lâmpadas, que podem estar apagadas, acesas com luz azul-clara ou acesas com luz azul-escura. Sob as lâmpadas estão três interruptores, A, B e C; cada um deles

Material com direitos autorais

muda o estado das lâmpadas para o *próximo* estado, na ordem mostrada na Figura 7. O interruptor A muda o estado das duas primeiras lâmpadas, o interruptor B muda todas as três lâmpadas e o interruptor C muda as últimas duas. Se todas as três lâmpadas estiverem inicialmente desligadas, é possível pressionar os interruptores de um jeito que as lâmpadas fiquem nesta ordem: desligada, azul-clara e azul-escura (como na Figura 8)?



Figura 7

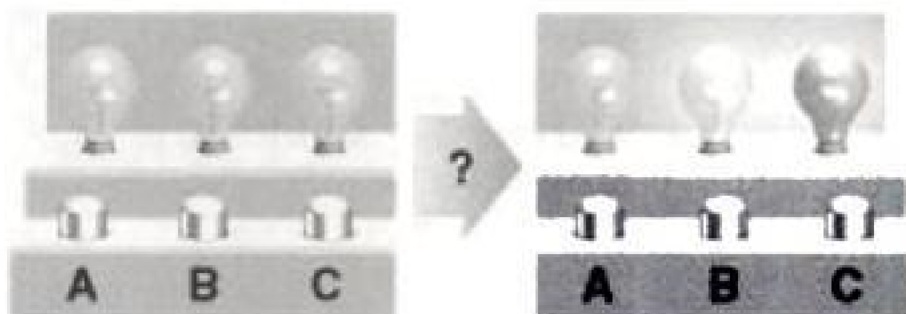


Figura 8

SOLUÇÃO: Enquanto o Exemplo 7 utilizou $\mathbb{Z}_2$, este exemplo claramente (está claro?) utiliza $\mathbb{Z}_3$. De acordo com isso, os interruptores correspondem aos vetores

$$\mathbf{a} = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \quad \mathbf{c} = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}$$

em $\mathbb{Z}_3^3$, e a configuração final que queremos obter é $\mathbf{t} = \begin{bmatrix} 0 \\ 1 \\ 2 \end{bmatrix}$. (Desligado é 0, azul-claro é 1 e azul-escuro é 2.)

Queremos encontrar escalares x_1, x_2, x_3 em $\mathbb{Z}_3$ tais que

$$x_1\mathbf{a} + x_2\mathbf{b} + x_3\mathbf{c} = \mathbf{t}$$

(onde x_i representa o número de vezes que o i-ésimo interruptor é pressionado. Essa equação faz surgir a matriz completa $[\mathbf{a}\ \mathbf{b}\ \mathbf{c} \mid \mathbf{t}]$, que se reduz em $\mathbb{Z}_3$ da seguinte maneira:

$$\left[\begin{array}{ccc|c} 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 2 \end{array}\right] \longrightarrow \left[\begin{array}{ccc|c} 1 & 0 & 0 & 2 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \end{array}\right]$$

Logo, há uma única solução para o sistema: $x_1 = 2$, $x_2 = 1$, $x_3 = 1$. Em outras palavras, devemos pressionar duas vezes o interruptor A e uma vez cada um dos outros dois interruptores. (Verifique isso.) ◆

◆ EXERCÍCIOS 2.5 ◆

Alocação de Recursos

1. Suponha que, no Exemplo 1, 400 unidades de alimento A, 600 unidades de alimento B e 600 unidades de alimento C sejam colocadas no tubo de ensaio a cada dia, e que os dados sobre o consumo diário de alimentos pelas bactérias (em unidades por dia) sejam como mostra a Tabela 3. Quantas bactérias de cada espécie podem coexistir no tubo de ensaio e consumir todo o alimento?

Tabela 3

	Bactéria da Espécie I	Bactéria da Espécie II	Bactéria da Espécie III
Alimento A	1	2	0
Alimento B	2	1	1
Alimento C	1	1	2

Material com direitos autorais

2. Suponha que, no Exemplo 1, 400 unidades de alimento A, 500 unidades de B e 600 unidades de C sejam colocadas no tubo a cada dia, e que os dados sobre o consumo diário de alimento pelas bactérias (em unidades por dia) sejam como mostra a Tabela 4. Quantas bactérias de cada espécie podem coexistir no tubo de ensaio e consumir todo o alimento?

Tabela 4

	Bactéria da Espécie I	Bactéria da Espécie II	Bactéria da Espécie III
Alimento A	1	2	0
Alimento B	2	1	3
Alimento C	1	1	1

3. Uma florista oferece três tamanhos de arranjos de flores com rosas, margaridas e crisântemos. Cada arranjo pequeno contém uma rosa, três margaridas e três crisântemos. Cada arranjo médio contém duas rosas, quatro margaridas e seis crisântemos. Cada arranjo grande contém quatro rosas, oito margaridas e seis crisântemos. Um dia, a florista notou que havia usado um total de 24 rosas, 50 margaridas e 48 crisântemos ao preparar as encomendas desses três tipos de arranjos. Quantos arranjos de cada tipo ela fez?

4. (a) Você tem no bolso algumas moedas de 5 centavos, de 10 centavos e de 25 centavos. Há 20 moedas no total e exatamente duas vezes mais moedas de 10 centavos do que de 5 centavos. O valor total das moedas é \$ 3,00. Encontre o número de moedas de cada tipo.

(b) Encontre *todas* as possíveis combinações de 20 moedas (de 5, 10 e 25 centavos) que totalizam \$ 3,00.

5. Um comerciante de café vende três misturas de grãos. Um pacote com a "mistura da casa" contém 300 gramas de café colombiano e 200 gramas de café tostado tipo francês. Um pacote com a "mistura especial" contém 200 gramas de café colombiano, 200 gramas de café queniano e 100 gramas de café tostado tipo francês. Um pacote com "mistura gourmet" contém 100 gramas de café colombiano, 200 gramas de café queniano e 200 gramas de café tostado tipo francês. O comerciante tem 30 quilos de café colombiano, 15 de café queniano e 25 de café tostado tipo francês. Se ele deseja utilizar todos os grãos de café, quantos pacotes de cada mistura deve preparar?

6. Refaça o Exercício 5, assumindo que a "mistura da casa" contém 300 gramas de café colombiano, 50 gramas de café queniano e 150 gramas de café tostado tipo francês, e a "mistura gourmet" contém 100 gramas de café colombiano, 350 gramas de café queniano e 50 gramas de café tostado tipo francês. Desta vez o comerciante tem à disposição 30 quilos de café colombiano, 15 de café queniano e 15 de café tostado tipo francês. Suponha que um pacote da "mistura da casa" dê um lucro de \$ 0,50, um pacote da mistura especial dê um lucro de \$ 1,50 e um pacote de "mistura gourmet" produza um lucro de \$ 2,00. Quantos pacotes de cada tipo o comerciante deve preparar se ele quer usar todo o estoque *e* maximizar seu lucro? Qual é o lucro máximo?

Balanceamento de Equações Químicas

Nos Exercícios de 7 a 14, faça o balanceamento da equação química para cada reação.


7. $FeS_2 + O_2 \longrightarrow Fe_2O_3 + SO_2$

8. $CO_2 + H_2O \longrightarrow C_6H_{12}O_6 + O_2$ (Essa reação ocorre quando uma planta verde converte dióxido de carbono e água em glicose e oxigênio durante a fotossíntese.)

9. $C_4H_{10} + O_2 \longrightarrow CO_2 + H_2O$ (Essa reação acontece quando o gás butano, C_4H_{10}, queima na presença de oxigênio para formar dióxido de carbono e água.)


10. $C_7H_6O_2 + O_2 \longrightarrow H_2O + CO_2$

11. $C_5H_{11}OH + O_2 \longrightarrow H_2O + CO_2$ (Essa equação representa a combustão do álcool pentanol.)


12. $HClO_4 + P_4O_{10} \longrightarrow H_3PO_4 + Cl_2O_7$


13. $Na_2CO_3 + C + N_2 \longrightarrow NaCN + CO$


14. $C_2H_2Cl_4 + Ca(OH)_2 \longrightarrow C_2HCl_3 + CaCl_2 + H_2O$

Análise de Redes

15. A Figura 9 mostra uma rede de canos de água com fluxo medido em litros por minuto.

(a) Monte e resolva um sistema de equações lineares para encontrar os fluxos possíveis.

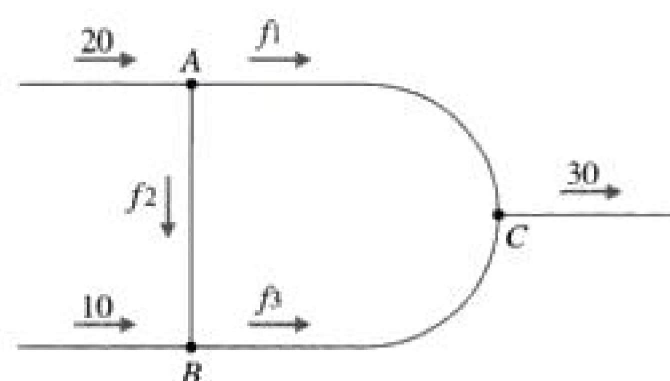


Figura 9

(b) Se o fluxo através de AB é restrito a 5 l/min, qual será o fluxo através dos outros dois ramos?

Material com direitos autorais

(c) Quais são os fluxos mínimo e máximo possíveis através de cada ramo?

(d) Até aqui, assumimos que o fluxo é sempre *positivo*. Qual seria o significado de um fluxo *negativo*, se o permitíssemos? Dê uma ilustração desse exemplo.

16. O coração do centro da cidade de Gotham consiste em ruas de mão única; o fluxo do tráfego é medido em cada cruzamento. Para os quarteirões mostrados na Figura 10, os números representam o número médio de veículos por minuto entrando e saindo dos cruzamentos A, B, C e D durante o horário comercial.

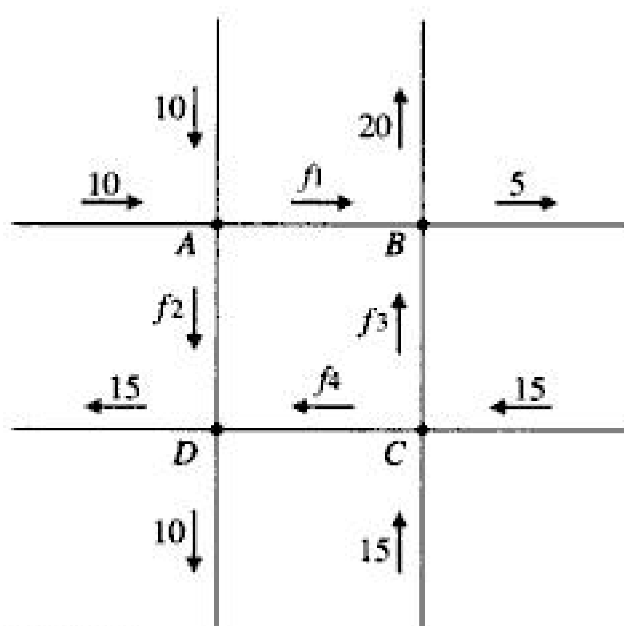


Figura 10

a) Monte e resolva um sistema de equações lineares para encontrar os possíveis fluxos.

(b) Se o tráfego em CD for regulado de modo que $f_4 = 10$ veículos por minuto, qual será o fluxo médio nas outras ruas?

(c) Quais são os fluxos mínimo e máximo possíveis em cada rua?

(d) De que forma o resultado mudaria se *todos* os sentidos do fluxo fossem trocados?

17. Uma rede de canais de irrigação é mostrada na Figura 11, com fluxo medido em milhares de litros por dia.

(a) Monte e resolva um sistema de equações lineares para encontrar os possíveis fluxos $f_1, \ldots, f_5$.

(b) Suponha que o canal DC esteja fechado. Qual intensidade de fluxo será necessário manter através de DB?

(c) Pela Figura 11, é claro que DB não pode ser fechado. (Por que não?) De que forma sua solução do item (a) mostra isso?

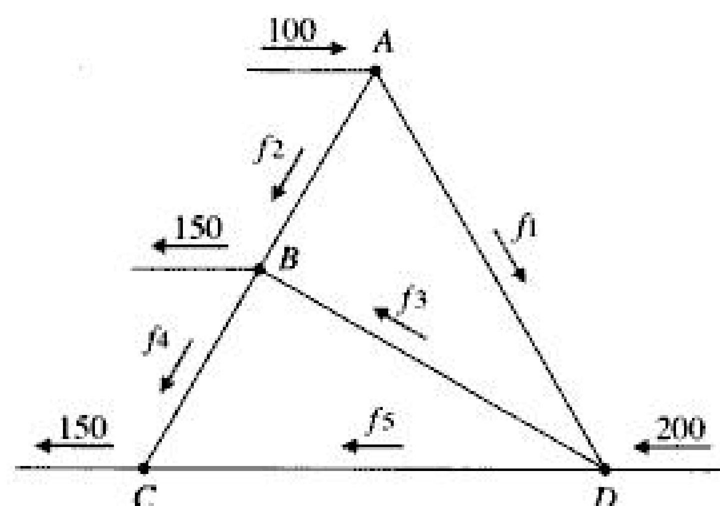


Figura 11

(d) Através de sua solução do item (a), determine os fluxos máximo e mínimo através de DB.

18. (a) Monte e resolva um sistema de equações lineares para encontrar os fluxos possíveis na rede mostrada na Figura 12.

(b) É possível ter $f_1 = 100$ e $f_6 = 150$? [Responda a essa pergunta primeiramente olhando sua solução do item (a) e depois de acordo com a Figura 12.]

(c) Se $f_4 = 0$, qual será a intensidade de fluxo em cada um dos outros ramos?

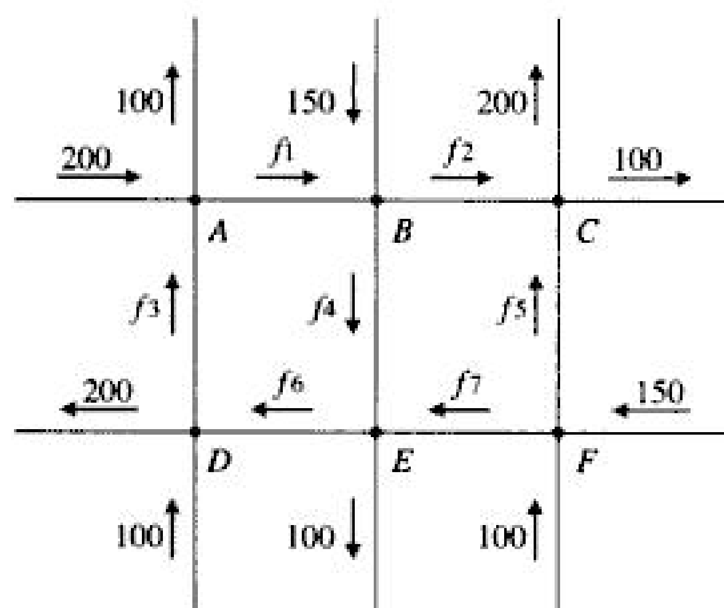


Figura 12

Circuitos Elétricos

Nos Exercícios 19 e 20, determine as correntes para cada circuito elétrico dado.

Material com direitos autorais



(b) Podemos pressionar os interruptores em alguma ordem de modo que apenas a segunda lâmpada fique acesa?

25. No Exemplo 7, descreva todas as possíveis configurações de lâmpadas que podem ser obtidas se começarmos com todas as lâmpadas apagadas.

26. (a) No Exemplo 8, suponha que todas as lâmpadas estejam inicialmente desligadas. Mostre que é possível pressionar os interruptores em alguma ordem de modo que as lâmpadas fiquem nesta ordem: desligada, azul-escura e azul-clara.

(b) Mostre que é possível pressionar os interruptores em alguma ordem de modo que as lâmpadas fiquem nesta ordem: azul-clara, desligada e azul-clara.

(c) Prove que *qualquer* configuração das três lâmpadas pode ser obtida.

27. Suponha que as lâmpadas do Exemplo 7 possam estar desligadas, azul-claras ou azul-escuras, e que os interruptores funcionem como no Exemplo 8 (isto é, os interruptores controlam as mesmas lâmpadas do Exemplo 7, mas por ciclos de cores, como no Exemplo 8). Mostre que é possível começar com todas as lâmpadas apagadas e pressionar os interruptores em alguma ordem de modo que as lâmpadas fiquem nesta ordem: azul-escura, azul-clara, azul-escura, azul-clara e azul-escura.

28. Para a mesma situação do Exercício 27, descreva todas as possíveis configurações das lâmpadas que podem ser obtidas, começando com todas as lâmpadas apagadas.

29. Nove quadrados, cada um ou branco ou preto, estão organizados em uma grade 3×3. A Figura 15 mostra uma configuração possível. Quando tocado, cada quadrado muda sua própria cor e a cor de alguns de seus vizinhos (preto → branco; branco → preto), conforme mostra a Figura 16. (Ao se tocar nos quadrados cujo número está circulado, mudam as cores dos quadrados marcados com *.) O objetivo do jogo é fazer todos os nove quadrados ficarem pretos. [Os Exercícios 29 e 30 são adaptados de quebra-cabeças que podem ser encontrados no CD-ROM de jogos interativos *The seventh guest* (Trilobyte Software/Virgin Games, 1992).]

1	2	3
4	5	6
7	8	9

Figura 15 O quebra-cabeças de nove quadrados

Figura 16 Mudanças de estado para o quebra-cabeças dos nove quadrados

(a) Se a configuração inicial é a mostrada na Figura 15, mostre que o jogo pode ser ganho e descreva uma seqüência de movimentos vencedora.

(b) Prove que o jogo pode sempre ser ganho, independentemente da configuração inicial.

30. Considere uma variação do quebra-cabeças dos nove quadrados. O jogo é como o descrito no Exercício 29, exceto que há três estados de cor possíveis para cada quadrado: branco, cinza ou preto. Os quadrados mudam de cor, como mostra a Figura 16, mas agora as mudanças de estado seguem o ciclo branco → cinza → preto → branco. Mostre como a configuração vencedora toda preta pode ser obtida a partir da configuração inicial mostrada na Figura 17.

1	2	3
4	5	6
7	8	9

Figura 17 O quebra-cabeças dos nove quadrados com mais estados

Miscelânea de Problemas

Nos Exercícios de 31 a 47, monte e resolva um sistema de equações lineares apropriado para responder às questões.

Material com direitos autorais

31. Graça é três vezes mais velha que Heitor, mas daqui a cinco anos terá o dobro da idade de Heitor. Qual a idade deles agora?

32. A soma das idades de Ana, Beto e Cristina é 60 anos. Ana é mais velha que Beto pelo mesmo número de anos que Beto é mais velho que Cristina. Quando Beto tiver a idade que Ana tem hoje, Ana terá três vezes a idade que Cristina tem hoje. Quais são suas idades?

Os dois problemas anteriores são típicos daqueles encontrados em livros populares de quebra-cabeças matemáticos. Entretanto, eles têm origem na antigüidade. Uma tábua de argila da Babilônia, de cerca de 300 a.C., contém o seguinte problema:

33. Há dois campos cuja área total é de 1.800 metros quadrados. Um campo produz grãos na proporção de $\frac{2}{3}$ de sacas por metro quadrado; o outro campo produz grãos na proporção de $\frac{1}{2}$ saca por metro quadrado. Se a produção total é de 1.100 sacas, qual é o tamanho de cada campo?

Há mais de 2.000 anos, os chineses desenvolveram métodos de resolução de sistemas de equações lineares, incluindo uma versão do método de eliminação de Gauss, que não se tornou conhecido na Europa até o século XIX. (Não há evidências de que Gauss conhecesse os métodos chineses quando desenvolveu aquele que chamamos atualmente de método de eliminação de Gauss. Porém, é claro que os chineses já conheciam a essência do método, embora eles não justificassem seu uso.) O seguinte problema foi extraído do texto chinês Jiuzhang suanshu (Nove capítulos em arte matemática)*, escrito durante a Dinastia de Han, cerca de 200 anos a.C.:*

34. Há três tipos de milho. Três feixes do primeiro tipo, dois do segundo e um do terceiro fazem 39 medidas. Dois feixes do primeiro tipo, três do segundo e um do terceiro fazem 34 medidas. Um feixe do primeiro tipo, dois do segundo e três do terceiro fazem 26 medidas. Quantas medidas de milho há em um feixe de cada tipo?

35. Descreva todos os possíveis valores de a, b, c e d que tornarão válidas as tabelas de adição seguintes. [Os problemas de 35 a 38 são baseados no artigo "An application of matrix theory", escrito por Paul Glaister em *The mathematics teacher*, 85 (1992), p. 220–223.]

(a)

+	a	b
c	2	3
d	4	5

(b)

+	a	b
c	3	6
d	4	5

36. Que condições sobre w, x, y e z irão garantir que possamos encontrar a, b, c e d de modo que a tabela de adição a seguir seja válida?

+	a	b
c	w	x
d	y	z

37. Descreva todos os valores de a, b, c, d, e e f que tornarão válidas as seguintes tabelas de adição:

(a)

+	a	b	c
d	3	2	1
e	5	4	3
f	4	3	1

(b)

+	a	b	c
d	1	2	3
e	3	4	5
f	4	5	6

38. Generalizando o Exercício 36, encontre condições sobre os dados de uma tabela de adição 3×3 que garantam que os valores de a, b, c, d, e e f possam sempre ser encontrados, como acima.

39. Sabemos, da geometria elementar, que existe uma única reta que passa por dois pontos distintos de um plano. É menos conhecido o fato de que existe uma única parábola que passa por quaisquer *três* pontos não colineares de um plano. Para cada conjunto de pontos a seguir, encontre uma parábola com equação da forma $y = ax^2 + bx + c$ que passe pelos pontos dados. (Esboce a parábola resultante para conferir a validade de sua resposta.)

(a) $(0, 1)$, $(-1, 4)$ e $(2, 1)$ (b) $(-3, 1)$, $(-2, 2)$ e $(-1, 5)$

40. Por três pontos não colineares também passa uma única circunferência. Encontre a circunferência (cuja equação geral é da forma $x^2 + y^2 + ax + by + c = 0$) que passa por cada conjunto de pontos dados no Exercício 39. (Para conferir a validade de sua resposta, encontre o centro e o raio de cada circunferência e desenhe um esboço.)

A adição de funções racionais (quociente de polinomiais), obtida através de uma escolha de um denominador comum, é feita de modo análogo à adição de números racionais. O processo reverso, de separar uma função racional escrevendo-a como uma soma de funções racionais mais simples, é útil em muitas áreas da matemática; por exemplo, aparece em cálculo diferencial e integral quando precisamos integrar uma função racional, e em matemática discreta, quando usamos funções geradoras para resolver relações de recorrência. A decomposição de uma função racional como soma de frações parciais leva a um sistema de equações lineares. Nos Exercícios 41 a 44, encontre a decomposição em frações parciais da forma dada. (As letras maiúsculas indicam constantes.)

41. $\dfrac{3x+1}{x^2+2x-3} = \dfrac{A}{x-1} + \dfrac{B}{x+3}$

42. $\dfrac{x^2-3x+3}{x^3+2x^2+x} = \dfrac{A}{x} + \dfrac{B}{x-1} + \dfrac{C}{(x-1)^2}$

43. $\dfrac{x-1}{(x+1)(x^2+1)(x^2+4)}$

$$= \frac{A}{x+1} + \frac{Bx+C}{x^2+1} + \frac{Dx+E}{x^2+4}$$

Material com direitos autorais

44. $$\frac{x^3 + x + 1}{x(x-1)(x^2+x+1)(x^2+1)^3} = \frac{A}{x} + \frac{B}{x-1} + \frac{Cx+D}{x^2+x+1} + \frac{Ex+F}{x^2+1} + \frac{Gx+H}{(x^2+1)^2} + \frac{Ix+J}{(x^2+1)^3}$$

Veja a seguir duas fórmulas úteis para as somas de potências de números naturais consecutivos:

$$1 + 2 + \cdots + n = \frac{n(n+1)}{2}$$

e $$1^2 + 2^2 + \cdots + n^2 = \frac{n(n+1)(2n+1)}{6}$$

A validade dessas fórmulas para todos os valores de $n \geq 1$ (ou mesmo $n \geq 0$) pode ser estabelecida usando-se indução matemática (veja o Apêndice B). Para fazer um "chute" com chances de acerto sobre como devem ser cada uma das fórmulas, observamos que é possível reescrevê-las como

$$\tfrac{1}{2}n^2 + \tfrac{1}{2}n \quad e \quad \tfrac{1}{3}n^3 + \tfrac{1}{2}n^2 + \tfrac{1}{6}n$$

respectivamente. Isso leva à conjectura de que a soma das p-ésimas potências dos primeiros n números naturais é um polinômio de grau $p + 1$ na variável n.

45. Assumindo que $1 + 2 + \cdots + n = an^2 + bn + c$, encontre a, b e c substituindo três valores de n e, portanto, obtendo um sistema de equações lineares em a, b e c.

46. Assuma que $1^2 + 2^2 + \cdots + n^2 = an^3 + bn^2 + cn + d$. Encontre a, b, c e d. (*Sugestão:* é correto usar $n = 0$. Como fica o lado esquerdo da equação nesse caso?)

47. Mostre que $1^3 + 2^3 + \cdots + n^3 = (n(n+1)/2)^2$.

CAS 2.6 Métodos Iterativos de Resolução de Sistemas Lineares

Os métodos diretos de resolução de sistemas lineares por meio de operações elementares com as linhas levam a soluções exatas em muitos casos, mas estão sujeitos a erros devidos ao arredondamento e outros fatores, como já vimos. A terceira estrada de nossas "três vias" nos leva a um caminho, na verdade, bastante diferente. Nesta seção, vamos explorar métodos que funcionam *iterativamente*, gerando seqüências sucessivas de vetores que se aproximam de uma solução de um sistema linear. Em muitos exemplos (tais como quando a matriz dos coeficientes é *esparsa* — isto é, contém muitos elementos nulos), os métodos iterativos podem ser mais rápidos e mais precisos que os métodos diretos. Além disso, métodos iterativos podem ser interrompidos sempre que a solução aproximada gerada por ele estiver suficientemente precisa. Os métodos iterativos, ainda, freqüentemente se *favorecem* da imprecisão: erros de arredondamento podem até acelerar sua convergência para uma solução.

> Carl Gustav Jacobi (1804–1851) foi um matemático alemão que fez contribuições importantes a muitas áreas da matemática e da física, incluindo geometria, teoria dos números, análise, mecânica e dinâmica dos fluidos. Embora boa parte de seu trabalho tenha sido em matemática aplicada, Jacobi acreditou na importância de fazer matemática pura. Bom professor, ele ocupou posições nas Universidades de Berlim e Königsberg e foi um dos mais famosos matemáticos na Europa.

Vamos explorar dois métodos iterativos para resolução de sistemas lineares: o ***método de Jacobi*** e um refinamento dele, o ***método de Gauss-Seidel***. Em todos os exemplos, consideraremos sistemas lineares com o mesmo número de incógnitas e de equações, e assumiremos que existe uma única solução. Nosso interesse é encontrar essa solução usando métodos iterativos.

EXEMPLO 1 Considere o sistema

$$\begin{aligned} 7x_1 - x_2 &= 5 \\ 3x_1 - 5x_2 &= -7 \end{aligned}$$

O método de Jacobi começa com a resolução da primeira equação para x_1 e da segunda equação para x_2, para obter

Material com direitos autorais

$$x_1 = \frac{5 + x_2}{7}$$

$$x_2 = \frac{7 + 3x_1}{5} \qquad (1)$$

Precisamos agora de uma ***aproximação inicial*** da solução. O que ocorre é que não importa qual seja a aproximação inicial, e assim podemos escolher $x_1 = 0$, $x_2 = 0$. Usamos esses valores nas equações (1) para conseguirmos novos valores de x_1 e x_2:

$$x_1 = \frac{5 + 0}{7} = \frac{5}{7} \approx 0{,}714$$

$$x_2 = \frac{7 + 3 \cdot 0}{5} = \frac{7}{5} = 1{,}400$$

Agora, substituímos esses valores em (1) e obtemos

$$x_1 = \frac{5 + 1{,}4}{7} \approx 0{,}914$$

$$x_2 = \frac{7 + 3 \cdot \frac{5}{7}}{5} \approx 1{,}829$$

Tabela 1

n	0	1	2	3	4	5	6
x_1	0	0,714	0,914	0,976	0,993	0,998	0,999
x_2	0	1,400	1,829	1,949	1,985	1,996	1,999

(valores escritos com três casas decimais). Repetimos esse processo (usando os valores anteriores de x_2 e de x_1 para obter os novos valores de x_1 e x_2), produzindo a seqüência de aproximações mostrada na Tabela 1.

Chamamos de ***iterados*** os vetores sucessivos $\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$. Por exemplo, quando $n = 4$, o quarto iterado é $\begin{bmatrix} 0{,}993 \\ 1{,}985 \end{bmatrix}$.

Podemos ver que os iterados, neste exemplo, estão se aproximando de $\begin{bmatrix} 1 \\ 2 \end{bmatrix}$, que é a solução exata do sistema dado. (Verifique.) Em um caso como esse, dizemos que o método de Jacobi ***converge***. ◆

O método de Jacobi calcula os iterados em um sistema de duas incógnitas de acordo com um padrão cruzado, conforme mostra a Tabela 2.

Tabela 2

n	0	1	2	3
x_1				
x_2				

Antes de considerarmos o método de Jacobi no caso geral, vamos olhar uma modificação dele que freqüentemente converge de modo mais rápido para a solução. O *método de Gauss-Seidel* é o mesmo que o

Material com direitos autorais

método de Jacobi, exceto pelo fato de usarmos cada valor novo *logo que possível.* Assim, em nosso exemplo, começamos calculando $x_1 = (5 + 0)/7 = \frac{5}{7} \approx 0{,}714$ como antes, mas agora usamos esse valor de x_1 para conseguir o próximo valor de x_2:

$$x_2 = \frac{7 + 3 \cdot \frac{5}{7}}{5} \approx 1{,}829$$

Usamos agora esse valor de x_2 para recalcular x_1, e assim por diante. Os iterados obtidos dessa maneira aparecem na Tabela 3.

Tabela 3

n	0	1	2	3	4	5
x_1	0	0,714	0,976	0,998	1,000	1,000
x_2	0	1,829	1,985	1,999	2,000	2,000

Observamos que o método de Gauss-Seidel convergiu para a solução mais rapidamente. Os iterados, desta vez, foram calculados de acordo com o padrão ziguezague mostrado na Tabela 4.

Tabela 4

n	0	1	2	3
x_1	↗	↓ ↗	↓ ↗	↓
x_2				

O método de Gauss-Seidel também tem uma interpretação geométrica interessante no caso de duas variáveis. Podemos pensar que x_1 e x_2 são as coordenadas de pontos do plano. Nosso ponto de partida é o ponto correspondente a nossa aproximação inicial, (0, 0). Nosso primeiro cálculo fornece $x_1 = \frac{5}{7}$, logo, nos movemos para o ponto $(\frac{5}{7}, 0) \approx (0{,}714, 0)$. Então calculamos $x_2 = \frac{64}{35} \approx 1{,}829$, que nos move para o ponto $(\frac{5}{7}, \frac{64}{35}) \approx (0{,}714, 1{,}829)$. Continuando dessa maneira, nossos cálculos pelo método de Gauss-Seidel farão surgir uma seqüência de pontos, cada um diferente do anterior em exatamente uma coordenada. Se fizermos os gráficos das retas $7x_1 - x_2 = 5$ e $3x_1 - 5x_2 = -7$, correspondentes às duas equações dadas, veremos que os pontos calcu-

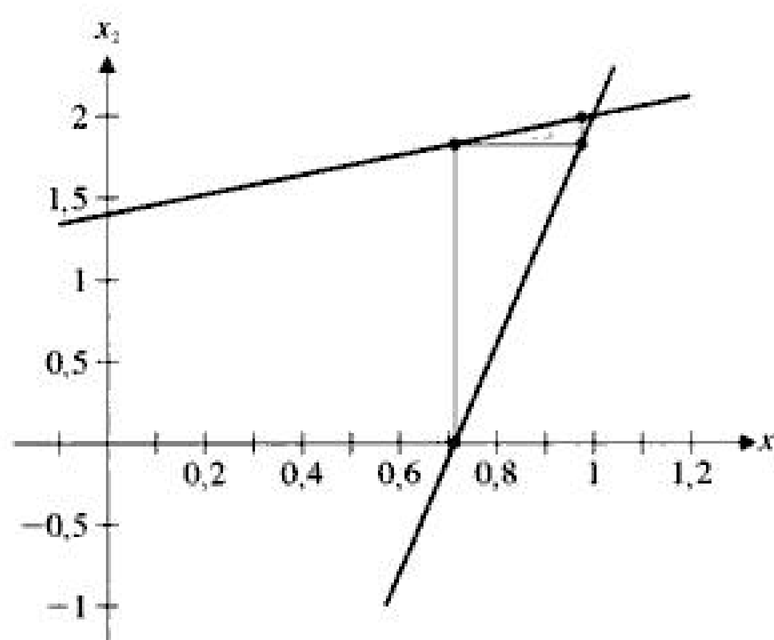


Figura 1 Iterados convergentes

Material com direitos autorais

O método de Gauss-Seidel tem esse nome em homenagem a C. F. Gauss e Philipp Ludwig von Seidel (1821–1896). Seidel trabalhou em análise, teoria da probabilidade, astronomia e óptica. Infelizmente, teve problemas na vista e se aposentou ainda jovem. O artigo no qual ele descreveu o método conhecido como Gauss-Seidel foi publicado em 1874. Gauss, ao que parece, não conhecia o método!

lados há pouco estão localizados sobre as duas retas alternadamente, como é mostrado na Figura 1. Além disso, eles se aproximam do ponto de interseção das retas, que corresponde à solução do sistema de equações. É esse o significado de *convergência*!

Os casos gerais dos dois métodos são análogos. Dado um sistema de n equações lineares em n incógnitas,

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= b_2 \\ &\vdots \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n &= b_n \end{aligned} \qquad (2)$$

resolvemos a primeira equação para x_1, a segunda para x_2 e assim por diante. Depois, começando com uma aproximação inicial qualquer, usamos essas novas equações para iterativamente atualizar cada variável. O método de Jacobi usa *todos* os valores obtidos na k-ésima iteração para calcular a $(k + 1)$-ésima iterada, enquanto o método de Gauss-Seidel sempre usa *os mais recentes* valores de cada variável em cada cálculo. O Exemplo 3 a seguir ilustra o método de Gauss-Seidel em um problema com três variáveis.

A essa altura, você deve ter algumas questões ou preocupações sobre esses métodos iterativos. (Você tem?) Várias vêm à mente: esses métodos sempre convergem? Se não, quando eles *realmente* convergem? *Se* convergem, é sempre para a solução? A resposta para a primeira questão é não, como mostra o Exemplo 2.

EXEMPLO 2 Aplique o método de Gauss-Seidel ao sistema

$$\begin{aligned} x_1 - x_2 &= 1 \\ 2x_1 + x_2 &= 5 \end{aligned}$$

com aproximação inicial $\begin{bmatrix} 0 \\ 0 \end{bmatrix}$.

SOLUÇÃO: Rearranjamos as equações para obter

$$\begin{aligned} x_1 &= 1 + x_2 \\ x_2 &= 5 - 2x_1 \end{aligned}$$

As primeiras iteradas são dadas na Tabela 5. (Confira.)

A solução real do sistema é $\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$. As iteradas na Tabela 5 não estão se aproximando desse ponto, como a Figura 2 mostra claramente através de um gráfico que exemplifica uma ***divergência***.

Tabela 5

n	0	1	2	3	4	5
x_1	0	1	4	−2	10	−14
x_2	0	3	−3	9	−15	33

Material com direitos autorais



◆ TEOREMA 2

Se o método de Jacobi ou o de Gauss-Seidel convergir em um sistema de n equações lineares em n incógnitas, então ele deve convergir para a solução do sistema.

DEMONSTRAÇÃO: Ilustraremos a idéia por trás da demonstração esboçando-a para o caso do método de Jacobi, usando o sistema de equações lineares do Exemplo 1. A prova geral é similar.

A convergência significa que, para alguma iteração, os valores das iteradas permanecem os mesmos. Isso significa que x_1 e x_2 convergem para r e s, respectivamente, como mostra a Tabela 6.

Devemos provar que $\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} r \\ s \end{bmatrix}$ é a solução do sistema de equações. Em outras palavras, na $(k + 1)$–ésima iteração, os valores de x_1 e x_2 devem permanecer os mesmos que na k-ésima iteração. Mas os cálculos dão $x_1 = (5 + x_2)/7 = (5 + s)/7$ e $x_2 = (7 + 3x_1)/5 = (7 + 3r)/5$. Portanto,

$$\frac{5+s}{7} = r \quad \text{e} \quad \frac{7+3r}{5} = s$$

Tabela 6

n	...	k	$k+1$	$k+2$	...
x_1	...	r	r	r	...
x_2	...	s	s	s	...

Rearranjando, vemos que

$$\begin{aligned} 7r - s &= 5 \\ 3r - 5s &= -7 \end{aligned}$$

Logo, $x_1 = r$, $x_2 = s$ satisfazem as equações originais, como desejado. ◆

Agora você deve estar curioso. Se os métodos iterativos nem sempre convergem para a solução, para que eles servem? Por que não usamos apenas o método de eliminação de Gauss? Primeiramente, vimos que o método de eliminação de Gauss é sensível a erros de arredondamento e pode nos levar a respostas não precisas ou até mesmo totalmente erradas. Além disso, mesmo quando o método de eliminação de Gauss não se desvia, não é possível melhorar uma solução que tenhamos encontrado. Por exemplo, se usamos o método de eliminação de Gauss para calcular uma solução com duas casas decimais de precisão, não há como obter a solução com quatro casas de precisão, exceto começando-se novamente e trabalhando com maior precisão.

Por outro lado, podemos alcançar precisão adicional com métodos iterativos simplesmente fazendo mais iterações. Para sistemas grandes, particularmente aqueles com matrizes dos coeficientes esparsas, métodos iterativos são muito mais rápidos que métodos diretos quando implementados em um computador. Em muitas aplicações, os sistemas que aparecem são estrita e diagonalmente dominantes, e, nesses casos, os métodos iterativos serão certamente convergentes. O próximo exemplo ilustra uma aplicação dessas.

EXEMPLO 3 Suponha que aqueçamos cada lado de uma placa de metal a uma temperatura constante, como mostra a Figura 3.

Com o passar do tempo, a temperatura em cada ponto do interior da placa alcançará um *equilíbrio*, e a seguinte propriedade poderá ser comprovada:

Material com direitos autorais

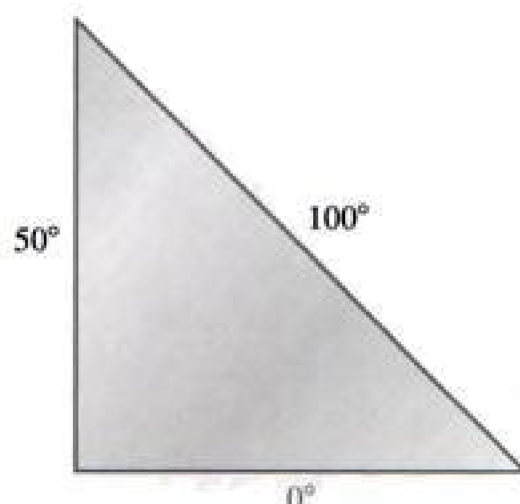


Figura 3 Uma placa de metal aquecida

A temperatura em cada ponto P interior em uma placa é a média das temperaturas na circunferência de qualquer círculo dentro da placa, centrado em P (Figura 4).

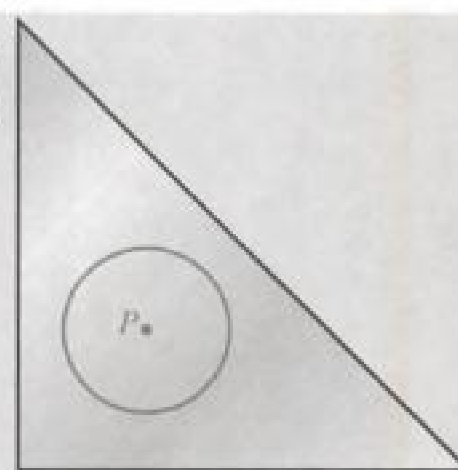


Figura 4

A aplicação dessa propriedade em um exemplo real requer técnicas de cálculo diferencial e integral. Como alternativa, podemos aproximar a situação colocando sobre a placa uma grade com um número finito de pontos de interseção, como mostra a Figura 5.

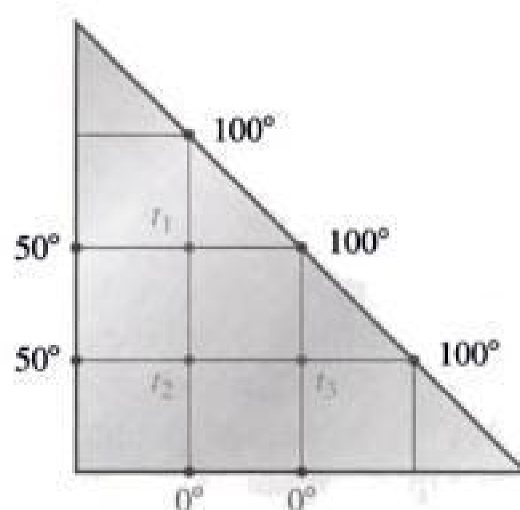


Figura 5 A versão discreta do problema da placa aquecida

O análogo discreto da propriedade da média que governa o equilíbrio de temperaturas é assim enunciado:

A temperatura em cada ponto interior P é a média das temperaturas nos pontos adjacentes a P.

No exemplo mostrado na Figura 5, há três pontos interiores, cada um adjacente a quatro outros pontos. Sejam t_1, t_2 e t_3 as temperaturas de equilíbrio dos pontos interiores, como mostrado. Então, pela propriedade da temperatura média, temos

Material com direitos autorais

$$t_1 = \frac{100 + 100 + t_2 + 50}{4}$$

$$t_2 = \frac{t_1 + t_3 + 0 + 50}{4} \qquad (3)$$

$$t_3 = \frac{100 + 100 + 0 + t_2}{4}$$

ou

$$\begin{aligned} 4t_1 - t_2 \qquad &= 250 \\ -t_1 + 4t_2 - t_3 &= 50 \\ - t_2 + 4t_3 &= 200 \end{aligned}$$

Note que esse sistema é estrita e diagonalmente dominante. Note também que as equações (3) são da forma requerida pelos métodos de iteração de Jacobi ou de Gauss-Seidel. Com uma aproximação inicial de $t_1 = 0$, $t_2 = 0$, $t_3 = 0$, o método de Gauss-Seidel dá as seguintes iterações:

Iteração 1:

$$t_1 = \frac{100 + 100 + 0 + 50}{4} = 62,5$$

$$t_2 = \frac{62,5 + 0 + 0 + 50}{4} = 28,125$$

$$t_3 = \frac{100 + 100 + 0 + 28,125}{4} = 57,031$$

Iteração 2:

$$t_1 = \frac{100 + 100 + 28,125 + 50}{4} = 69,531$$

$$t_2 = \frac{69,531 + 57,031 + 0 + 50}{4} = 44,141$$

$$t_3 = \frac{100 + 100 + 0 + 44,141}{4} = 61,035$$

Continuando, encontramos as iterações listadas na Tabela 7. Trabalhamos com cinco algarismos significativos e paramos quando duas iterações sucessivas têm valores que diferem em menos de 0,001 nas três variáveis.
Assim, as temperaturas de equilíbrio nos pontos interiores são (dentro da faixa de erro de 0,001) $t_1 = 74,108$, $t_2 = 46,430$ e $t_3 = 61,607$. (Confira esses cálculos.)

Tabela 7

n	0	1	2	3	···	7	8
t_1	0	62,500	69,531	73,535	···	74,107	74,107
t_2	0	28,125	44,141	46,143	···	46,429	46,429
t_3	0	57,031	61,035	61,536	···	61,607	61,607

◆

Usando uma malha mais fina (com mais pontos), podemos obter informações tão precisas quanto necessário sobre as temperaturas de equilíbrio nos vários pontos do plano.

Material com direitos autorais





(d) Resolva o sistema de equações exatamente e compare suas respostas.

28. Uma formiga está parada sobre o ponto A de uma reta. Ela anda metade do caminho para o ponto B e se vira. Em seguida, anda metade do caminho de volta para o ponto A, vira-se novamente e anda metade do caminho para o ponto B. Continua a fazer isso indefinidamente. Suponha que o ponto A seja 0 e o ponto B, 1. O passeio da formiga é feito de uma seqüência de segmentos de reta que se sobrepõem. Sejam x_1 o registro das posições dos extremos à esquerda desses segmentos e x_2 o das posições à direita. (Assim, começamos com $x_1 = 0$ e $x_2 = \frac{1}{2}$. Então, temos $x_1 = \frac{1}{4}$ e $x_2 = \frac{1}{2}$, e assim por diante.) A Figura 7 mostra o começo do passeio da formiga.

(a) Faça uma tabela com os seis primeiros valores de $[x_1, x_2]$ e marque os pontos correspondentes em um sistema de coordenadas x_1–x_2.

(b) Encontre duas equações lineares, da forma $x_2 = ax_1 + b$ e $x_1 = cx_2 + d$, que determinem os novos valores das extremidades em cada iteração. Desenhe as retas correspondentes em seu sistema de coordenadas e mostre que esse diagrama resultaria da aplicação do método de Gauss-Seidel ao sistema de equações lineares que você encontrou. (Seu diagrama deve ser parecido com a Figura 1 na página 116.)

(c) Trocando para representação decimal, continue aplicando o método de Gauss-Seidel para aproximar os valores para os quais x_1 e x_2 estão convergindo, com precisão de 0,001.

(e) Resolva o sistema de equações exatamente e compare suas respostas. Interprete seus resultados.

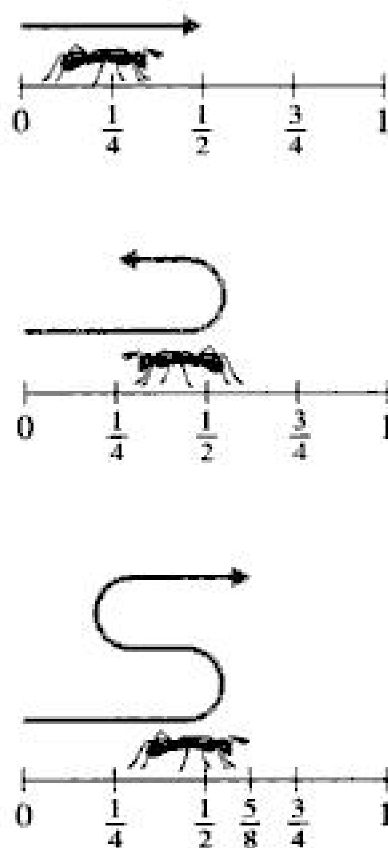


Figura 7 O passeio da formiga

Material com direitos autorais

3

Matrizes

Nós [Halmos e Kaplansky] compartilhamos uma filosofia sobre álgebra linear: pensamos em base-livre, escrevemos em base-livre, mas, quando as dificuldades surgem, fechamos a porta de nossos escritórios e calculamos com matrizes ferozmente.

– Irving Kaplansky
Em *Paul Halmos: Celebrating 50 years of mathematics*
J. H. Ewing e F. W. Gehring, Eds.
Springer-Verlag, 1991, p. 88

3.1 Introdução: Matrizes em Ação

Neste capítulo, estudaremos as matrizes por si mesmas. Já usamos matrizes — na forma de matrizes completas — para gravar informações sobre sistemas de equações lineares e para facilitar a resolução deles. Agora, veremos que as matrizes têm propriedades algébricas próprias, que nos capacitam a calcular seguindo as regras da álgebra de matrizes. Além disso, observaremos que matrizes não são objetos estáticos, que gravam informações e dados; mais do que isso, elas representam certos tipos de funções e "agem" em vetores, transformando-os em outros vetores. Essas "transformações por meio de matrizes" começarão a representar um papel-chave em nosso estudo da álgebra linear e darão uma nova visão do que aprendemos sobre vetores e sobre sistemas de equações lineares. Mais ainda, matrizes surgem em muitas formas além da forma de matrizes completas; investigaremos algumas das muitas aplicações das matrizes no final deste capítulo.

Nesta seção, apresentaremos alguns exemplos simples para ilustrar como matrizes podem transformar vetores. No processo, você terá uma primeira noção da "aritmética matricial".

Considere as equações

$$\begin{aligned} y_1 &= x_1 + 2x_2 \\ y_2 &= \phantom{x_1 + {}} 3x_2 \end{aligned} \tag{1}$$

Podemos ver que essas equações descrevem a transformação do vetor $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$ no vetor $\mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$. Se denotarmos a matriz de coeficientes do lado direito da equação por F, então $F = \begin{bmatrix} 1 & 2 \\ 0 & 3 \end{bmatrix}$, e poderemos reescrever a transformação como

Material com direitos autorais

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} 1 & 2 \\ 0 & 3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

ou, mais sucintamente, $\mathbf{y} = F\mathbf{x}$. [Pense nessa expressão como análoga à notação de função $y = f(x)$, com a qual você está familiarizado: $\mathbf{x}$ é aqui a "variável" independente, $\mathbf{y}$ é a "variável" dependente e F é o nome da "função".]

Assim, se $\mathbf{x} = \begin{bmatrix} -2 \\ 1 \end{bmatrix}$, as equações (1) resultam em

$$\begin{aligned} y_1 &= -2 + 2 \cdot 1 = 0 \\ y_2 &= 3 \cdot 1 = 3 \end{aligned} \quad \text{ou} \quad \mathbf{y} = \begin{bmatrix} 0 \\ 3 \end{bmatrix}$$

Podemos escrever essa expressão como $\begin{bmatrix} 0 \\ 3 \end{bmatrix} = \begin{bmatrix} 1 & 2 \\ 0 & 3 \end{bmatrix} \begin{bmatrix} -2 \\ 1 \end{bmatrix}$.

Problema 1: Calcule $F\mathbf{x}$ para os seguintes vetores $\mathbf{x}$:

(a) $\mathbf{x} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ (b) $\mathbf{x} = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$ (c) $\mathbf{x} = \begin{bmatrix} -1 \\ -1 \end{bmatrix}$ (d) $\mathbf{x} = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$

Problema 2: Os quatro vetores $\mathbf{x}$ do Problema 1 representam os quatro vértices de um quadrado no plano x_1x_2. Desenhe esse quadrado e nomeie os vértices de A', B', C' e D', correspondendo às partes (a), (b), (c), e (d) do Problema 1.

Em um outro sistema de coordenadas (chamando os eixos de y_1 e y_2), marque os quatro pontos determinados por $F\mathbf{x}$ no Problema 1. Nomeie esses pontos como A', B', C' e D'. Façamos a suposição (razoável) de que o segmento de reta $\overline{AB}$ seja transformado no segmento de reta $\overline{A'B'}$, o mesmo acontecendo com os outros três lados do quadrado $ABCD$. Que figura geométrica é representada por $A'B'C'D'$?

Problema 3: O centro do quadrado $ABCD$ é a origem $\mathbf{0} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$. Qual é o centro de A', B', C' e D'? Que cálculo algébrico confirma isso?

Agora, considere as equações

$$\begin{aligned} z_1 &= y_1 - y_2 \\ z_2 &= -2y_1 \end{aligned} \qquad (2)$$

que transformam um vetor $\mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$ no vetor $\mathbf{z} = \begin{bmatrix} z_1 \\ z_2 \end{bmatrix}$. Podemos abreviar essa transformação como $\mathbf{z} = G\mathbf{y}$, em que

$$G = \begin{bmatrix} 1 & -1 \\ -2 & 0 \end{bmatrix}$$

Problema 4: Descobriremos como G transforma a figura $A'B'C'D'$. Calcule $G\mathbf{y}$ para cada um dos quatro vetores $\mathbf{y}$ que você calculou no Problema 1 [isto é, calcule $\mathbf{z} = G(F\mathbf{x})$. Você pode reconhecer essa expressão como análoga à composição de funções que já lhe é familiar]. Chame os novos pontos de A'', B'', C'' e D'' e desenhe a figura $A''B''C''D''$ nos eixos coordenados z_1z_2.

Problema 5: Substituindo as equações (1) nas equações (2), obtemos equações para z_1 e z_2 em termos de x_1 e x_2. Se denotarmos a matriz dessas equações como H, teremos $\mathbf{z} = H\mathbf{x}$. Como também temos que $\mathbf{z} = GF\mathbf{x}$, é razoável escrever

$$H = GF$$

Você consegue perceber como os elementos de H estão relacionados com os elementos de F e de G?

Material com direitos autorais

Problema 6: Vamos resolver de trás para a frente o processo que acabamos de ver. Primeiro, transforme o quadrado $ABCD$, usando G, para obter a figura $A^*B^*C^*D^*$. Depois, transforme a figura resultante usando F para obter $A^{**}B^{**}C^{**}D^{**}$. [*Observação*: não se preocupe aqui com as "variáveis" **x**, **y** e **z**. Simplesmente substitua as coordenadas de A, B, C e D nas equações (2) e então substitua os resultados nas equações (1).] $A^{**}B^{**}C^{**}D^{**}$ e $A''B''C''D''$ são iguais? O que isso lhe mostra sobre a ordem em que aplicamos as transformações F e G?

Problema 7: Repita o Problema 5 com as matrizes genéricas

$$F = \begin{bmatrix} f_{11} & f_{12} \\ f_{21} & f_{22} \end{bmatrix}, \quad G = \begin{bmatrix} g_{11} & g_{12} \\ g_{21} & g_{22} \end{bmatrix} \quad \text{e} \quad H = \begin{bmatrix} h_{11} & h_{12} \\ h_{21} & h_{22} \end{bmatrix}$$

Se as equações (1) e (2) tiverem coeficientes como especificado por F e G, ache os elementos de H em termos dos elementos de F e G. O resultado será uma fórmula para o "produto" $H = GF$.

Problema 8: Repita os Problemas de 1 a 6 com as matrizes a seguir. (Sua fórmula obtida no Problema 7 pode ajudar a agilizar os cálculos algébricos.) Note as semelhanças e diferenças que você considerar significantes.

$$\text{(a) } F = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}, G = \begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix} \qquad \text{(b) } F = \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix}, G = \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix}$$

$$\text{(c) } F = \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix}, G = \begin{bmatrix} 2 & -1 \\ -1 & 1 \end{bmatrix} \qquad \text{(d) } F = \begin{bmatrix} 1 & -2 \\ -2 & 4 \end{bmatrix}, G = \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix}$$

3.2 Operações com Matrizes

Apesar de já termos encontrado matrizes, começamos dando uma definição formal e registrando alguns fatos para referência futura.

Definição Uma ***matriz*** é uma tabela retangular de números chamados de ***elementos*** ou ***termos*** da matriz.[1]

Estes são exemplos de matrizes:

$$\begin{bmatrix} 1 & 2 \\ 0 & 3 \end{bmatrix}, \quad \begin{bmatrix} \sqrt{5} & -1 & 0 \\ 2 & \pi & \frac{1}{2} \end{bmatrix}, \quad \begin{bmatrix} 2 \\ 4 \\ 17 \end{bmatrix}, \quad [1 \quad 1 \quad 1 \quad 1], \quad \begin{bmatrix} 5{,}1 & 1{,}2 & -1 \\ 6{,}9 & 0 & 4{,}4 \\ -7{,}3 & 9 & 8{,}5 \end{bmatrix}, \quad [7]$$

A ***ordem*** de uma matriz descreve o número de linhas e colunas que ela tem. Uma matriz é chamada $m \times n$ (pronuncia-se "m por n") quando tem m linhas e n colunas. Conseqüentemente, os exemplos que acabamos de ver são matrizes de ordem 2×2, 2×3, 3×1, 1×4, 3×3 e 1×1, respectivamente. Uma matriz $1\times m$ é chamada de ***matriz linha*** (ou ***vetor linha***), e uma matriz $n\times1$ é chamada de ***matriz coluna*** (ou ***vetor coluna***).[2]

[1] Embora esses números sejam normalmente escolhidos do conjunto dos números reais, $\mathbb{R}$, eles também podem ser tomados do conjunto dos números complexos, $\mathbb{C}$, ou dos inteiros módulo p, $\mathbb{Z}_p$, onde p é primo.

[2] Tecnicamente, existe diferença entre matrizes linha/coluna e vetores, mas não focaremos essa distinção. Entretanto, distinguiremos matrizes linha/vetores de matrizes coluna/vetores. Essa distinção é importante — pelo menos — para os cálculos algébricos, como demonstraremos.

Material com direitos autorais







You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.





Tabela 3

	Sam	Téo
Ana	$ 2,80	$ 3,80
Beto	$ 4,10	$ 4,00

Assim, o produto da matriz DP nos mostra quanto a compra de cada um em cada banca irá custar (Tabela 3). ◆

EXEMPLO 8 Considere o sistema linear

$$\begin{aligned} x_1 - 2x_2 + 3x_3 &= 5 \\ -x_1 + 3x_2 + x_3 &= 1 \\ 2x_1 - x_2 + 4x_3 &= 14 \end{aligned} \qquad (1)$$

Observe que o lado esquerdo dessa igualdade surge do produto de matrizes

$$\begin{bmatrix} 1 & -2 & 3 \\ -1 & 3 & 1 \\ 2 & -1 & 4 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

portanto, o sistema (1) pode ser escrito da forma

$$\begin{bmatrix} 1 & -2 & 3 \\ -1 & 3 & 1 \\ 2 & -1 & 4 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 5 \\ 1 \\ 14 \end{bmatrix}$$

ou $A\mathbf{x} = \mathbf{b}$, em que A é a matriz de coeficientes, $\mathbf{x}$ é o vetor (coluna) das variáveis e $\mathbf{b}$ é o vetor (coluna) dos termos constantes. ◆

Você não deve ter dificuldades em perceber que *todos* os sistemas lineares podem ser escritos na forma $A\mathbf{x} = \mathbf{b}$. Na verdade, a notação $[A \mid \mathbf{b}]$ para a matriz completa de um sistema linear é simplesmente uma abreviação para a equação matricial $A\mathbf{x} = \mathbf{b}$. Essa forma se revelará como um meio tremendamente útil de expressar um sistema de equações lineares, e nós a exploraremos freqüentemente daqui por diante.

Combinando essa idéia com o Teorema 1 da Seção 2.4, vemos que $A\mathbf{x} = \mathbf{b}$ tem solução se, e somente se, $\mathbf{b}$ for uma combinação linear das colunas de A.

Também provaremos ser muito útil um outro fato sobre as matrizes: a multiplicação de uma matriz por um vetor canônico unitário pode ser usada para "selecionar" ou "reproduzir" uma coluna ou uma linha de uma matriz. Seja $A = \begin{bmatrix} 4 & 2 & 1 \\ 0 & 5 & -1 \end{bmatrix}$ e considere os produtos $A\mathbf{e}_3$ e $\mathbf{e}_2A$, com os vetores unitários $\mathbf{e}_3$ e $\mathbf{e}_2$ escolhidos de forma que os produtos façam sentido. Assim,

$$A\mathbf{e}_3 = \begin{bmatrix} 4 & 2 & 1 \\ 0 & 5 & -1 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ -1 \end{bmatrix} \quad \text{e} \quad \mathbf{e}_2A = [0 \quad 1] \begin{bmatrix} 4 & 2 & 1 \\ 0 & 5 & -1 \end{bmatrix} = [0 \quad 5 \quad -1]$$

Observe que $A\mathbf{e}_3$ nos dá a terceira coluna de A, e $\mathbf{e}_2A$ nos dá a segunda linha de A. Enunciamos o resultado geral como um teorema.

Material com direitos autorais

◆ TEOREMA 1

Sejam A uma matriz $m \times n$, $\mathbf{e}_i$ um vetor canônico unitário $1 \times m$ e $\mathbf{e}_j$ um vetor canônico unitário $n \times 1$. Então:

$\mathbf{e}_i A$ é a i-ésima linha da matriz A e
$A\mathbf{e}_j$ é a j-ésima coluna da matriz A.

DEMONSTRAÇÃO: Provamos (b) e deixamos para provar (a) no Exercício 41. Se $\mathbf{a}_1, \ldots, \mathbf{a}_n$ são as colunas de A, o produto $A\mathbf{e}_j$ pode ser escrito na forma

$$A\mathbf{e}_j = 0\mathbf{a}_1 + 0\mathbf{a}_2 + \ldots + 1\mathbf{a}_j + \ldots + 0\mathbf{a}_n = \mathbf{a}_j$$

Poderíamos também provar (b) pelo cálculo:

$$A\mathbf{e}_j = \begin{bmatrix} a_{11} & \cdots & a_{1j} & \cdots & a_{1n} \\ a_{21} & \cdots & a_{2j} & \cdots & a_{2n} \\ \vdots & & \vdots & & \\ a_{m1} & \cdots & a_{mj} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{bmatrix} = \begin{bmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{mj} \end{bmatrix}$$

pois o 1 em $\mathbf{e}_j$ é o j-ésimo elemento. ◆

Decomposição de Matrizes

Muitas vezes é conveniente olhar uma matriz como composta por um número de ***submatrizes*** menores. Traçando-se linhas horizontais e verticais em uma matriz, podemos ***decompô-la*** em ***blocos***. Existe uma maneira natural de decompor matrizes em blocos, particularmente aquelas provenientes de certas aplicações. Por exemplo, considere a matriz

$$A = \begin{bmatrix} 1 & 0 & 0 & 2 & -1 \\ 0 & 1 & 0 & 1 & 3 \\ 0 & 0 & 1 & 4 & 0 \\ 0 & 0 & 0 & 1 & 7 \\ 0 & 0 & 0 & 7 & 2 \end{bmatrix}$$

Parece natural decompor A na forma

$$\left[\begin{array}{ccc|cc} 1 & 0 & 0 & 2 & -1 \\ 0 & 1 & 0 & 1 & 3 \\ 0 & 0 & 1 & 4 & 0 \\ \hline 0 & 0 & 0 & 1 & 7 \\ 0 & 0 & 0 & 7 & 2 \end{array}\right] = \begin{bmatrix} I & B \\ O & C \end{bmatrix}$$

em que I é a matriz identidade 3×3, B é 3×2, O é a matriz nula 2×3 e C é 2×2. Dessa forma, podemos ver A como uma matriz 2×2 cujos elementos são matrizes.

Quando se multiplicam matrizes, há freqüentemente vantagem ao vê-las como matrizes decompostas em blocos. Isso não só revela muitas vezes estruturas escondidas como também quase sempre acelera o cálculo, especialmente quando as matrizes são grandes e têm muitos blocos de zeros. Fica claro que a multiplicação de matrizes decompostas em blocos é igual à multiplicação comum de matrizes.

Material com direitos autorais

Começamos considerando alguns casos especiais de matrizes decompostas em blocos. Cada um deles mostra um modo diferente de ver o produto de duas matrizes.

Suponha que A seja uma matriz $m \times n$ e B, $n \times r$. O produto AB, portanto, existe. Se decompusermos B em termos de seus vetores coluna, como $B = [\mathbf{b}_1 \mid \mathbf{b}_2 \mid \cdots \mid \mathbf{b}_r]$, então

$$AB = A[\mathbf{b}_1 \mid \mathbf{b}_2 \mid \cdots \mid \mathbf{b}_r] = [A\mathbf{b}_1 \mid A\mathbf{b}_2 \mid \cdots \mid A\mathbf{b}_r]$$

Esse resultado é uma conseqüência imediata da definição de multiplicação de matriz. A forma da direita é chamada de ***representação matriz-coluna*** do produto.

EXEMPLO 9 Se

$$A = \begin{bmatrix} 1 & 3 & 2 \\ 0 & -1 & 1 \end{bmatrix} \quad \text{e} \quad B = \begin{bmatrix} 4 & -1 \\ 1 & 2 \\ 3 & 0 \end{bmatrix}$$

então

$$A\mathbf{b}_1 = \begin{bmatrix} 1 & 3 & 2 \\ 0 & -1 & 1 \end{bmatrix}\begin{bmatrix} 4 \\ 1 \\ 3 \end{bmatrix} = \begin{bmatrix} 13 \\ 2 \end{bmatrix} \quad \text{e} \quad A\mathbf{b}_2 = \begin{bmatrix} 1 & 3 & 2 \\ 0 & -1 & 1 \end{bmatrix}\begin{bmatrix} -1 \\ 2 \\ 0 \end{bmatrix} = \begin{bmatrix} 5 \\ -2 \end{bmatrix}$$

Portanto, $AB = [A\mathbf{b}_1 \mid A\mathbf{b}_2] = \begin{bmatrix} 13 & 5 \\ 2 & -2 \end{bmatrix}$. (Verifique fazendo a multiplicação usual.) ◆

Observação: • Note que a representação matriz-coluna de AB nos permite escrever cada coluna de AB como uma combinação das colunas de A, tendo os elementos de B como coeficientes. Por exemplo,

$$\begin{bmatrix} 13 \\ 2 \end{bmatrix} = \begin{bmatrix} 1 & 3 & 2 \\ 0 & -1 & 1 \end{bmatrix}\begin{bmatrix} 4 \\ 1 \\ 3 \end{bmatrix} = 4\begin{bmatrix} 1 \\ 0 \end{bmatrix} + \begin{bmatrix} 3 \\ -1 \end{bmatrix} + 3\begin{bmatrix} 2 \\ 1 \end{bmatrix}$$

(Veja os Exercícios 23 e 26.)

Suponha que A seja uma matriz $m \times n$ e B, $n \times r$. O produto AB, portanto, existe. Se decompusermos A em termos de seus vetores linha, como

$$A = \begin{bmatrix} \mathbf{A}_1 \\ \hline \mathbf{A}_2 \\ \hline \vdots \\ \hline \mathbf{A}_m \end{bmatrix}$$

então

$$AB = \begin{bmatrix} \mathbf{A}_1 \\ \hline \mathbf{A}_2 \\ \hline \vdots \\ \hline \mathbf{A}_m \end{bmatrix} B = \begin{bmatrix} \mathbf{A}_1 B \\ \hline \mathbf{A}_2 B \\ \hline \vdots \\ \hline \mathbf{A}_m B \end{bmatrix}$$

Material com direitos autorais





(Quando alguns dos blocos são matrizes nulas, ou matrizes identidade, como no nosso caso, esses cálculos podem ser feitos muito rapidamente.) Os cálculos para os outros cinco blocos de AB são semelhantes. Verifique que o resultado é

$$\left[\begin{array}{cc|cc|c} 6 & 2 & 1 & 2 & 2 \\ 0 & 5 & 2 & 1 & 12 \\ 5 & -5 & 3 & 3 & 9 \\ \hline 1 & 7 & 0 & 0 & 23 \\ 7 & 2 & 0 & 0 & 20 \end{array}\right]$$

(Observe que o bloco do lado superior esquerdo é o resultado do nosso último cálculo.) Verifique que você obtém a mesma resposta multiplicando A por B da forma usual. ◆

Potências de Matriz

Quando A e B forem duas matrizes $n \times n$, o produto delas também será uma matriz $n \times n$. Um caso especial ocorre quando $A = B$. Faz sentido definir $A^2 = AA$ e, em geral, definir A^k como

$$A^k = \underbrace{AA \cdots A}_{k \text{ fatores}}$$

sendo k um inteiro positivo. Assim, $A^1 = A$, e é conveniente definir $A^0 = I_n$.

Antes de fazer muitas suposições, precisamos nos perguntar com que extensão as potências de matrizes se comportam como as potências dos números reais. As propriedades a seguir originam-se imediatamente das definições que acabamos de dar e são matrizes análogas das propriedades correspondentes para potências de números reais.

> Se A é uma matriz quadrada e r e s são inteiros não negativos, então
> 1. $A^rA^s = A^{r+s}$
> 2. $(A^r)^s = A^{rs}$

Na Seção 3.4, estenderemos a definição e as propriedades para incluir potências com números inteiros negativos.

EXEMPLO 13 (a) Se $A = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$, então

$$A^2 = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 2 & 2 \\ 2 & 2 \end{bmatrix}, \quad A^3 = A^2A = \begin{bmatrix} 2 & 2 \\ 2 & 2 \end{bmatrix}\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 4 & 4 \\ 4 & 4 \end{bmatrix}$$

e, em geral,

$$A^n = \begin{bmatrix} 2^{n-1} & 2^{n-1} \\ 2^{n-1} & 2^{n-1} \end{bmatrix} \text{ para todo } n \geq 1$$

Essa afirmação pode ser provada por indução matemática, já que ela é dada para uma coleção *infinita* de afirmações, uma para cada número natural n. (O Apêndice B mostra uma breve revisão de indução matemática.) O primeiro passo é provar que a fórmula vale para $n = 1$. Nesse caso,

Material com direitos autorais

$$A^1 = \begin{bmatrix} 2^{1-1} & 2^{1-1} \\ 2^{1-1} & 2^{1-1} \end{bmatrix} = \begin{bmatrix} 2^0 & 2^0 \\ 2^0 & 2^0 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} = A$$

como queríamos.

A hipótese de indução é assumir que

$$A^k = \begin{bmatrix} 2^{k-1} & 2^{k-1} \\ 2^{k-1} & 2^{k-1} \end{bmatrix}$$

para algum número inteiro $k \geq 1$. O passo de indução é para provar que a fórmula vale para $n = k + 1$. Usando a definição de potência de matriz e a hipótese de indução, calculamos

$$\begin{aligned} A^{k+1} = A^k A &= \begin{bmatrix} 2^{k-1} & 2^{k-1} \\ 2^{k-1} & 2^{k-1} \end{bmatrix}\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \\ &= \begin{bmatrix} 2^{k-1} + 2^{k-1} & 2^{k-1} + 2^{k-1} \\ 2^{k-1} + 2^{k-1} & 2^{k-1} + 2^{k-1} \end{bmatrix} \\ &= \begin{bmatrix} 2^k & 2^k \\ 2^k & 2^k \end{bmatrix} \\ &= \begin{bmatrix} 2^{(k+1)-1} & 2^{(k+1)-1} \\ 2^{(k+1)-1} & 2^{(k+1)-1} \end{bmatrix} \end{aligned}$$

Assim, a fórmula vale para todo $n \geq 1$, pelo princípio da indução matemática.

(b) Se $B = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$, então $B^2 = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix}$. Continuando, encontramos

$$B^3 = B^2B = \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix}\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$$

e

$$B^4 = B^3B = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

Assim, $B^5 = B$, e a seqüência de potências de B se repete em um ciclo de quatro:

$$\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}, \ldots$$

A Transposta de uma Matriz

Até agora, todas as operações com matrizes que definimos são análogas a operações com números reais, embora elas nem sempre se comportem da mesma maneira. A próxima operação não tem essa analogia.

A ***transposta*** de uma matriz A $m \times n$ é uma matriz $n \times m$ A^T obtida pela permutação das linhas de A pelas suas colunas — isto é, a i-ésima coluna de A^T é a i-ésima linha de A para todo i.

EXEMPLO 14 Sejam

$$A = \begin{bmatrix} 1 & 3 & 2 \\ 5 & 0 & 1 \end{bmatrix}, \quad B = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \quad \text{e} \quad C = [5 \quad -1 \quad 2]$$

Material com direitos autorais

Então, suas transpostas são

$$A^T = \begin{bmatrix} 1 & 5 \\ 3 & 0 \\ 2 & 1 \end{bmatrix}, \quad B^T = \begin{bmatrix} a & c \\ b & d \end{bmatrix} \quad \text{e} \quad C^T = \begin{bmatrix} 5 \\ -1 \\ 2 \end{bmatrix}$$ ◆

A transposta é algumas vezes usada como uma definição alternativa para o produto escalar de dois vetores, em termos de multiplicação de matriz. Se

$$\mathbf{u} = \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix} \quad \text{e} \quad \mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix}$$

então

$$\begin{aligned} \mathbf{u} \cdot \mathbf{v} &= u_1v_1 + u_2v_2 + \cdots + u_nv_n \\ &= [u_1 \quad u_2 \quad \cdots \quad u_n] \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix} \\ &= \mathbf{u}^T\mathbf{v} \end{aligned}$$

Uma definição alternativa útil para a transposta é dada elemento a elemento:

$$(A^T)_{ij} = A_{ji} \quad \text{para todos } i \text{ e } j$$

Em palavras, o elemento na linha i e coluna j de A^T é igual ao elemento na linha j e coluna i de A.

A transposta é também usada para definir um tipo muito importante de matriz quadrada: a matriz simétrica. Uma matriz é ***simétrica*** quando é igual à sua transposta.

EXEMPLO 15 Sejam

$$A = \begin{bmatrix} 1 & 3 & 2 \\ 3 & 5 & 0 \\ 2 & 0 & 4 \end{bmatrix} \quad \text{e} \quad B = \begin{bmatrix} 1 & 2 \\ -1 & 3 \end{bmatrix}$$

Então, A é simétrica, pois $A^T = A$; entretanto, B não é simétrica, já que $B^T = \begin{bmatrix} 1 & -1 \\ 2 & 3 \end{bmatrix} \neq B$. ◆

Uma matriz simétrica tem a propriedade de ser a sua própria "imagem refletida" através da diagonal principal. A Figura 1 ilustra essa propriedade para uma matriz 3×3. As formas geométricas correspondentes representam elementos iguais; os elementos da diagonal (representados em linha tracejada) são arbitrários.

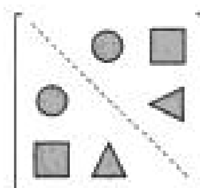


Figura 1 Uma matriz simétrica

Material com direitos autorais

Uma definição de matriz simétrica elemento a elemento também é útil. Ela é simplesmente a descrição algébrica da propriedade da "reflexão".

Uma matriz quadrada A é simétrica se, e somente se, $A_{ij} = A_{ji}$ para todos i e j.

◆ EXERCÍCIOS 3.2 ◆

Sejam

$$A = \begin{bmatrix} 3 & 0 \\ -1 & 5 \end{bmatrix}, B = \begin{bmatrix} 4 & -2 & 1 \\ 0 & 2 & 3 \end{bmatrix}, C = \begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{bmatrix},$$

$$D = \begin{bmatrix} 0 & -3 \\ -2 & 1 \end{bmatrix}, E = [4 \quad 2], F = \begin{bmatrix} -1 \\ 2 \end{bmatrix}$$

Nos Exercícios de 1 a 16, calcule a matriz indicada (se possível).

1. $A + 2D$

2. $3D - 2A$

3. $B - C$

4. $B - C^T$

5. AB

6. BD

7. $D + BC$

8. B^TB

9. $E(AF)$

10. $F(DF)$

11. FE

12. EF

13. $B^TC^T - (CB)^T$

14. $DA - AD$

15. A^3

16. $(I_2 - D)^2$

17. Dê um exemplo de uma matriz A 2×2 não nula tal que $A^2 = O$.

18. Seja $A = \begin{bmatrix} 2 & 1 \\ 6 & 3 \end{bmatrix}$. Encontre as matrizes 2×2 B e C tais que $AB = AC$, mas $B \neq C$.

19. Uma fábrica produz três produtos (banheiras, pias e tanques) e os envia para armazenamento em dois depósitos. O número de unidades enviadas de cada produto para cada depósito é dado pela matriz

$$A = \begin{bmatrix} 200 & 75 \\ 150 & 100 \\ 100 & 125 \end{bmatrix}$$

(em que a_{ij} é o número de unidades enviadas do produto i para o depósito j, e os produtos são colocados em ordem alfabética). O custo de remessa de uma unidade de cada produto, por caminhão, é: \$ 1,50 por banheira, \$ 1,00 por pia e \$ 2,00 por tanque. Os custos unitários correspondentes ao envio por trem são: \$ 1,75, \$ 1,50 e \$ 1,00. Organize esses custos em uma matriz B e use essa matriz para mostrar como a fábrica pode comparar os custos de remessa — por caminhão e por trem — de seus produtos para cada um dos dois depósitos.

20. Em relação ao Exercício 19, suponha que o custo unitário de distribuição dos produtos para as lojas seja o mesmo para todos os produtos, mas que varie dependendo do depósito por causa das distâncias envolvidas. Custa \$ 0,75 para distribuir uma unidade do depósito 1 e \$ 1,00 para distribuir uma unidade do depósito 2. Organize esses custos em uma matriz C e use multiplicação de matrizes para calcular o custo total de distribuição de cada produto.

Nos Exercícios 21 e 22, escreva o sistema de equações lineares dado na forma de equação matricial, $A\mathbf{x} = \mathbf{b}$.

21.
$$\begin{aligned} x_1 - 2x_2 + 3x_3 &= 0 \\ 2x_1 + x_2 - 5x_3 &= 4 \end{aligned}$$

22.
$$\begin{aligned} -x_1 \quad\quad + 2x_3 &= 1 \\ x_1 - x_2 \quad\quad &= -2 \\ x_2 + x_3 &= -1 \end{aligned}$$

Nos Exercícios de 23 a 28, sejam

$$A = \begin{bmatrix} 1 & 0 & -2 \\ -3 & 1 & 1 \\ 2 & 0 & -1 \end{bmatrix}$$

e

$$B = \begin{bmatrix} 2 & 3 & 0 \\ 1 & -1 & 1 \\ -1 & 6 & 4 \end{bmatrix}$$

23. Use a representação matriz-coluna do produto para escrever cada coluna de AB como combinação linear das colunas de A.

24. Use a representação matriz-linha do produto para escrever cada linha de AB como combinação linear das linhas de B.

25. Calcule o produto externo da expansão de AB.

26. Use a representação matriz-coluna do produto para escrever cada coluna de BA como combinação linear das colunas de B.

27. Use a representação matriz-linha do produto para escrever cada linha de BA como combinação linear das linhas de A.

Material com direitos autorais



Propriedades da Adição e da Multiplicação por Escalar

Todas as propriedades algébricas da adição e da multiplicação por escalar dos vetores (Teorema 1 da Seção 1.2) valem para matrizes. Resumimos essas propriedades no próximo teorema.

◆ **TEOREMA 1 Propriedades Algébricas da Adição de Matrizes e da Multiplicação por Escalar**

Sejam A, B e C matrizes de mesma ordem, e c e d, escalares. Então:

a. $A + B = B + A$ — Comutativa
b. $(A + B) + C = A + (B + C)$ — Associativa
c. $A + O = A$
d. $A + (-A) = O$
e. $c(A + B) = cA + cB$ — Distributiva
f. $(c + d)A = cA + dA$ — Distributiva
g. $c(dA) = (cd)A$
h. $1A = A$

As demonstrações dessas propriedades são análogas às demonstrações correspondentes das propriedades dos vetores. Da mesma forma, os comentários que seguem o Teorema 1 da Seção 1.2 são igualmente válidos aqui, e não deveremos ter dificuldades para usar essas propriedades para fazer manipulações algébricas com matrizes. (Consulte o Exemplo 5 da Seção 1.2 e veja os Exercícios 17 e 18 no final desta seção.)

A propriedade associativa nos permite combinar a multiplicação por escalar e a adição sem fazer uso de parênteses. Se A, B e C são matrizes de mesma ordem, então

$$(2A + 3B) - C = 2A + (3B - C)$$

e assim podemos simplesmente escrever $2A + 3B - C$. Mais geralmente, se $A_1, A_2, \ldots, A_k$ são matrizes de mesma ordem e $c_1, c_2, \ldots, c_k$ são escalares, fazemos a ***combinação linear***

$$c_1A_1 + c_2A_2 + \ldots + c_kA_k$$

Chamaremos $c_1, c_2, \ldots, c_k$ de ***coeficientes*** da combinação linear. Vamos agora elaborar e responder questões sobre combinações lineares de matrizes.

EXEMPLO 1 Sejam $A_1 = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$, $A_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ e $A_3 = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$.

(a) $B = \begin{bmatrix} 1 & 4 \\ 2 & 1 \end{bmatrix}$ é combinação linear de A_1, A_2 e A_3?

(b) $C = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$ é combinação linear de A_1, A_2 e A_3?

SOLUÇÃO:

(a) Queremos encontrar escalares c_1, c_2 e c_3 tais que $c_1A_1 + c_2A_2 + c_3A_3 = B$. Assim,

$$c_1\begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} + c_2\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + c_3\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 4 \\ 2 & 1 \end{bmatrix}$$

O lado esquerdo dessa equação pode ser escrito na forma:

$$\begin{bmatrix} c_2 + c_3 & c_1 + c_3 \\ -c_1 + c_3 & c_2 + c_3 \end{bmatrix}$$

Material com direitos autorais



Encontraremos tais paralelos repetidamente daqui por diante. No Capítulo 6, eles serão explorados com mais detalhes.

Podemos definir o ***conjunto gerado*** por um conjunto de matrizes como o conjunto de todas as combinações lineares dessas matrizes.

EXEMPLO 2 Descreva o conjunto gerado pelas matrizes A_1, A_2 e A_3 do Exemplo 1.

SOLUÇÃO: Um maneira de fazer isso é simplesmente escrever uma combinação linear genérica de A_1, A_2 e A_3. Assim,

$$c_1A_1 + c_2A_2 + c_3A_3 = c_1\begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} + c_2\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + c_3\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$$
$$= \begin{bmatrix} c_2 + c_3 & c_1 + c_3 \\ -c_1 + c_3 & c_2 + c_3 \end{bmatrix}$$

(que é análogo à representação paramétrica de um plano). Porém, suponha que queiramos saber quando a matriz $\begin{bmatrix} w & x \\ y & z \end{bmatrix}$ estará no ger (A_1, A_2, A_3). Da representação que acabamos de ver, sabemos que esse fato se dará quando

$$\begin{bmatrix} c_2 + c_3 & c_1 + c_3 \\ -c_1 + c_3 & c_2 + c_3 \end{bmatrix} = \begin{bmatrix} w & x \\ y & z \end{bmatrix}$$

para alguma escolha de escalares c_1, c_2, c_3. Isso dará surgimento a um sistema de equações lineares cujo lado esquerdo será exatamente o mesmo do Exemplo 1, mas cujo lado direito será genérico. A matriz completa desse sistema será

$$\left[\begin{array}{rrr|c} 0 & 1 & 1 & w \\ 1 & 0 & 1 & x \\ -1 & 0 & 1 & y \\ 0 & 1 & 1 & z \end{array}\right]$$

e o escalonamento por linhas produzirá

$$\left[\begin{array}{rrr|c} 0 & 1 & 1 & w \\ 1 & 0 & 1 & x \\ -1 & 0 & 1 & y \\ 0 & 1 & 1 & z \end{array}\right] \longrightarrow \left[\begin{array}{ccc|c} 1 & 0 & 0 & \frac{1}{2}x - \frac{1}{2}y \\ 0 & 1 & 0 & -\frac{1}{2}x - \frac{1}{2}y + w \\ 0 & 0 & 1 & \frac{1}{2}x + \frac{1}{2}y \\ 0 & 0 & 0 & w - z \end{array}\right]$$

(Verifique isso cuidadosamente.) A única restrição se refere à última linha, onde claramente devemos ter $w - z = 0$ para que haja solução. Portanto, o conjunto gerado por A_1, A_2 e A_3 consiste em todas as matrizes $\begin{bmatrix} w & x \\ y & z \end{bmatrix}$ para as quais $w = z$. Ou seja, o ger$(A_1, A_2, A_3) = \left\{\begin{bmatrix} w & x \\ y & w \end{bmatrix}\right\}$. ◆

Nota: Se tivéssemos observado isso *antes*, olhando o Exemplo 1, teríamos visto imediatamente que $B = \begin{bmatrix} 1 & 4 \\ 2 & 1 \end{bmatrix}$ é uma combinação linear de A_1, A_2 e A_3, pois B tem a forma necessária (tome $w = 1$, $x = 4$ e $y = 2$), mas $C = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$ não pode ser uma combinação linear de A_1, A_2 e A_3, pois não tem a forma apropriada ($1 \neq 4$).

Material com direitos autorais

A independência linear também faz sentido para as matrizes. Dizemos que matrizes $A_1, A_2, \ldots, A_k$ de mesma ordem são ***linearmente independentes*** se a única solução da equação

$$c_1A_1 + c_2A_2 + \ldots + c_kA_k = O \tag{1}$$

é a solução trivial: $c_1 = c_2 = \ldots = c_k = 0$. Se existirem coeficientes não triviais que satisfaçam (1), então $A_1, A_2, \ldots, A_k$ serão chamadas de ***linearmente dependentes***.

EXEMPLO 3 Determine se as matrizes A_1, A_2 e A_3 do Exemplo 1 são linearmente independentes.

SOLUÇÃO: Queremos resolver a equação $c_1A_1 + c_2A_2 + c_3A_3 = O$. Escrevendo as matrizes, temos

$$c_1\begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} + c_2\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + c_3\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$

Desta vez, obtemos um sistema linear *homogêneo* cujo lado esquerdo é o mesmo que o dos exemplos 1 e 2. (Você já está conseguindo vislumbrar um padrão?) A matriz completa escalonada por linhas dá

$$\left[\begin{array}{ccc|c} 0 & 1 & 1 & 0 \\ 1 & 0 & 1 & 0 \\ -1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 \end{array}\right] \longrightarrow \left[\begin{array}{ccc|c} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right]$$

Portanto, $c_1 = c_2 = c_3 = 0$, e concluímos que as matrizes A_1, A_2 e A_3 são linearmente independentes. ◆

Propriedades da Multiplicação de Matrizes

Sempre que encontramos uma operação nova, tal como a multiplicação de matrizes, devemos ser cuidadosos para não presumir muito sobre ela. Seria bom se a multiplicação de matrizes se comportasse como a multiplicação de números reais. Embora em muitos aspectos isso aconteça, existem algumas diferenças significativas.

EXEMPLO 4 Considere as matrizes

$$A = \begin{bmatrix} 2 & 4 \\ -1 & -2 \end{bmatrix} \quad \text{e} \quad B = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}$$

Multiplicando-as, temos

$$AB = \begin{bmatrix} 2 & 4 \\ -1 & -2 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 6 & 4 \\ -3 & -2 \end{bmatrix} \quad \text{e} \quad BA = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}\begin{bmatrix} 2 & 4 \\ -1 & -2 \end{bmatrix} = \begin{bmatrix} 2 & 4 \\ 1 & 2 \end{bmatrix}$$

Portanto, $AB \neq BA$. Assim, em contraste com a multiplicação de números reais, a multiplicação de matrizes *não é comutativa* — a ordem dos fatores no produto é importante!

É fácil verificar que $A^2 = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$ (faça-o!). Assim, para matrizes, a equação $A^2 = O$ não implica que $A = O$ (diferentemente da situação para números reais, em que a equação $x^2 = 0$ tem somente $x = 0$ como solução). ◆

Apesar de coisas obscuras poderem aparecer depois do último exemplo, a situação não é tão ruim — você precisa somente acostumar-se a trabalhar com as matrizes e lembrar-se constantemente de que elas não são números. O próximo teorema resume as propriedades principais de multiplicação de matrizes.

Material com direitos autorais





assim,

$$BB^T = \begin{bmatrix} 4 & -1 & 0 \\ 2 & 3 & 1 \end{bmatrix} \begin{bmatrix} 4 & 2 \\ -1 & 3 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 17 & 5 \\ 5 & 14 \end{bmatrix}$$

e

$$B^TB = \begin{bmatrix} 4 & 2 \\ -1 & 3 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 4 & -1 & 0 \\ 2 & 3 & 1 \end{bmatrix} = \begin{bmatrix} 20 & 2 & 2 \\ 2 & 10 & 3 \\ 2 & 3 & 1 \end{bmatrix}$$

Portanto, BB^T e B^TB são simétricas, embora B nem mesmo seja quadrada! (Verifique que AA^T e A^TA são também simétricas.) ◆

O próximo teorema diz que os resultados do Exemplo 6 são em geral verdadeiros.

◆ TEOREMA 4

a. Se A é uma matriz quadrada, então $A + A^T$ é uma matriz simétrica.
b. Para toda matriz A, AA^T e A^TA são matrizes simétricas.

DEMONSTRAÇÃO: Provamos (a) e deixamos para provar (b) no Exercício 34. Simplesmente observamos que

$$(A + A^T)^T = A^T + (A^T)^T = A^T + A = A + A^T$$

(usando propriedades da transposta e a comutativa da adição de matrizes). Assim, $A + A^T$ é igual à sua transposta, por isso, por definição, é simétrica. ◆

◆ EXERCÍCIOS 3.3◆

Nos Exercícios de 1 a 4, ache X, dadas $A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$ *e* $B = \begin{bmatrix} -1 & 0 \\ 1 & 1 \end{bmatrix}$.

1. $X - 2A + 3B = O$

2. $2X = A - B$

3. $2(A + 2B) = 3X$

4. $2(A - B + X) = 3(X - A)$

Nos Exercícios de 5 a 8, escreva B como combinação linear das outras matrizes, se possível.

5. $B = \begin{bmatrix} 2 & 5 \\ 0 & 3 \end{bmatrix}, A_1 = \begin{bmatrix} 1 & 2 \\ -1 & 1 \end{bmatrix}, A_2 = \begin{bmatrix} 0 & 1 \\ 2 & 1 \end{bmatrix}$

6. $B = \begin{bmatrix} 2 & 3 \\ -4 & 2 \end{bmatrix}, A_1 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, A_2 = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix},$
$A_3 = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$

7. $B = \begin{bmatrix} 3 & 1 & 1 \\ 0 & 1 & 0 \end{bmatrix}, A_1 = \begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & 0 \end{bmatrix},$
$A_2 = \begin{bmatrix} -1 & 2 & 0 \\ 0 & 1 & 0 \end{bmatrix}, A_3 = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 0 & 0 \end{bmatrix}$

8. $B = \begin{bmatrix} 2 & -2 & 3 \\ 0 & 0 & -2 \\ 0 & 0 & 2 \end{bmatrix}, A_1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix},$
$A_2 = \begin{bmatrix} 0 & 1 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}, A_3 = \begin{bmatrix} -1 & 0 & -1 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{bmatrix},$
$A_4 = \begin{bmatrix} 1 & -1 & 1 \\ 0 & -1 & -1 \\ 0 & 0 & 1 \end{bmatrix}$

Nos Exercícios de 9 a 12, encontre a expressão geral do conjunto gerado pelas matrizes dadas, como no Exemplo 2.

Material com direitos autorais

9. $\text{ger}(A_1, A_2)$ do Exercício 5

10. $\text{ger}(A_1, A_2, A_3)$ do Exercício 6

11. $\text{ger}(A_1, A_2, A_3)$ do Exercício 7

12. $\text{ger}(A_1, A_2, A_3, A_4)$ do Exercício 8

Nos Exercícios de 13 a 16, determine se as matrizes dadas são linearmente independentes.

13. $\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}, \begin{bmatrix} 4 & 3 \\ 2 & 1 \end{bmatrix}$

14. $\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}, \begin{bmatrix} 2 & 1 \\ -1 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 2 \\ 4 & 3 \end{bmatrix}$

15. $\begin{bmatrix} 0 & 1 \\ 5 & 2 \\ -1 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 2 & 3 \\ 1 & 1 \end{bmatrix}, \begin{bmatrix} -2 & -1 \\ 0 & 1 \\ 0 & 2 \end{bmatrix}, \begin{bmatrix} -1 & -3 \\ 1 & 9 \\ 4 & 5 \end{bmatrix}$

16. $\begin{bmatrix} 1 & -1 & 0 \\ 0 & 2 & 0 \\ 0 & 2 & 6 \end{bmatrix}, \begin{bmatrix} 2 & 1 & 0 \\ 0 & 3 & 0 \\ 0 & 4 & 9 \end{bmatrix}, \begin{bmatrix} 1 & 2 & 0 \\ 0 & 1 & 0 \\ 0 & 3 & 5 \end{bmatrix},$

$\begin{bmatrix} -1 & 1 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -4 \end{bmatrix}$

17. Prove o Teorema 1(a) – (d).

18. Prove o Teorema 1(e) – (h).

19. Prove o Teorema 2(c).

20. Prove o Teorema 2(d).

21. Prove a parte do Teorema 2(e) que não foi provada no texto.

22. Prove que, para matrizes quadradas A e B, $AB = BA$ se, e somente se, $(A - B)(A + B) = A^2 - B^2$.

Nos Exercícios de 23 a 25, se $B = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$, *encontre condições em a, b, c e d para que* $AB = BA$.

23. $A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$

24. $A = \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}$

25. $A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$

26. Encontre condições em a, b, c e d para que $B = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ comute com $\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$ e $\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$.

27. Encontre condições em a, b, c e d para que $B = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ comute com todas as matrizes 2×2.

28. Prove que, se AB e BA estiverem definidas, AB e BA são matrizes quadradas.

Uma matriz quadrada é chamada de ***triangular superior*** *quando todos os elementos abaixo da diagonal principal são zeros. Assim, a forma de uma matriz triangular superior é*

$$\begin{bmatrix} * & * & \cdots & * & * \\ 0 & * & \cdots & * & * \\ 0 & 0 & \ddots & \vdots & \vdots \\ \vdots & \vdots & & * & * \\ 0 & 0 & \cdots & 0 & * \end{bmatrix}$$

*em que os elementos marcadas por * são arbitrários. Uma definição mais formal dessa matriz* $A = [a_{ij}]$ é que $a_{ij} = 0$ *se* $i > j$.

29. Prove que o produto de duas matrizes triangulares superiores $n\times n$ é uma matriz triangular superior.

30. Prove o Teorema 3(a) – (c).

31. Prove o Teorema 3(e).

32. Usando indução, prove que $(A_1 + A_2 + \cdots + A_n)^T = A_1^T + A_2^T + \cdots + A_n^T$ para todo $n \geq 1$.

33. Usando indução, prove que $(A_1 A_2 \cdots A_n)^T = A_n^T \cdots A_2^T A_n^T$ para todo $n \geq 1$.

34. Prove o Teorema 4(b).

35. (a) Prove que, se A e B são matrizes $n\times n$ simétricas, $A + B$ também é.

(b) Prove que, se A é uma matriz $n\times n$ simétrica, kA também é para todo escalar k.

36. (a) Dê um exemplo para mostrar que, se A e B são matrizes $n\times n$ simétricas, AB não é necessariamente simétrica.

(b) Prove que, se A e B são matrizes $n\times n$ simétricas, AB é simétrica se, e somente se, $AB = BA$.

Uma matriz quadrada é chamada de ***anti-simétrica*** *se* $A^T = -A$.

37. Quais das seguintes matrizes são anti-simétricas?

(a) $\begin{bmatrix} 1 & 2 \\ -2 & 3 \end{bmatrix}$ (b) $\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$

(c) $\begin{bmatrix} 0 & 3 & -1 \\ -3 & 0 & 2 \\ 1 & -2 & 0 \end{bmatrix}$ (d) $\begin{bmatrix} 0 & 1 & 2 \\ -1 & 0 & 5 \\ 2 & 5 & 0 \end{bmatrix}$

38. Dê uma definição de matriz anti-simétrica elemento a elemento.

39. Prove que a diagonal principal de uma matriz anti-simétrica é formada inteiramente por zeros.

40. Prove que, se A e B são matrizes $n\times n$ anti-simétricas, $A + B$ também é.

41. Se A e B são matrizes 2×2 anti-simétricas, sob que condições AB é anti-simétrica?

42. Prove que, se A é uma matriz $n\times n$, $A - A^T$ é anti-simétrica.

Material com direitos autorais



EXEMPLO 2 Mostre que as seguintes matrizes não são invertíveis:

(a) $O = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$ (b) $B = \begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix}$

SOLUÇÃO:
(a) É fácil ver que a matriz nula O não tem inversa. Se tivesse, existiria uma matriz O' tal que $OO' = I = O'O$. Entretanto, o produto da matriz nula por qualquer outra matriz é a matriz nula, e assim OO' nunca será igual à matriz identidade I. (Observe que essa demonstração não faz referência à ordem das matrizes, por isso ela é verdadeira para matrizes $n \times n$ em geral.)

(b) Suponha que B tenha uma inversa $B' = \begin{bmatrix} w & x \\ y & z \end{bmatrix}$. A equação $BB' = I$ nos dá

$$\begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix}\begin{bmatrix} w & x \\ y & z \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

da qual obtemos as equações

$$\begin{aligned} w \quad\quad + 2y \quad\quad &= 1 \\ x \quad\quad + 2z &= 0 \\ 2w \quad\quad + 4y \quad\quad &= 0 \\ 2x \quad\quad + 4z &= 1 \end{aligned}$$

Subtraindo duas vezes a primeira equação da terceira, obtemos $0 = -2$, que obviamente é um absurdo, por isso não existe solução. (O escalonamento por linhas fornece o mesmo resultado, mas ele não é realmente necessário nesse caso.) Deduzimos que a matriz B' não existe, ou seja, que B não é invertível. (De fato, ela não tem nem mesmo uma inversa em um dos lados, imagine então nos dois!) ◆

Observações:

- Embora tenhamos visto que a multiplicação de matrizes não é em geral comutativa, A' (se existir) deve satisfazer $A'A = AA'$.
- Os exemplos que acabamos de ver levantaram duas questões que iremos responder nesta seção:
  (1) Como podemos saber quando uma matriz tem uma inversa?
  (2) Se uma matriz tem uma inversa, como podemos encontrá-la?
- Não questionamos ainda a possibilidade de uma matriz A ter mais de uma inversa. O próximo teorema nos assegura que isso não pode acontecer.

◆ TEOREMA 1

Se A é uma matriz invertível, então a sua inversa é única.

DEMONSTRAÇÃO: Em matemática, um modo padrão de mostrar que existe apenas uma de alguma coisa é mostrar que não pode existir mais de uma. Assim, suponha que A tenha duas inversas — digamos, A' e A''. Então:

$$AA' = I = A'A \quad \text{e} \quad AA'' = I = A''A$$

Assim,

$$A' = A'I = A'(AA'') = (A'A)A'' = IA'' = A''$$

Portanto, $A' = A''$, e a inversa é única. ◆

Material com direitos autorais

Graças a esse teorema, podemos agora nos referir *à* inversa de uma matriz invertível. De agora em diante, quando A for invertível, denotaremos a sua inversa (única) por A^{-1} (chamada "inversa de A").

Aviso: Não pense em escrever $A^{-1} = \frac{1}{A}$! Não existe a operação "divisão por matriz". Mesmo que ela existisse, como poderíamos dividir o *escalar* 1 pela *matriz* A? Se você alguma vez se sentir tentado a "dividir" por uma matriz, o que você realmente quer fazer é multiplicar pela sua inversa.

Podemos agora completar a analogia que mencionamos no começo desta seção.

◆ **TEOREMA 2**

Se A é uma matriz invertível $n \times n$, o sistema de equações lineares dado por $A\mathbf{x} = \mathbf{b}$ tem uma única solução $\mathbf{x} = A^{-1}\mathbf{b}$ para todo $\mathbf{b}$ em $\mathbb{R}^n$.

DEMONSTRAÇÃO: O Teorema 2 essencialmente formaliza a observação que fizemos no começo desta seção. Voltaremos a ela mais cuidadosamente desta vez. Precisamos provar duas coisas: que $A\mathbf{x} = \mathbf{b}$ *tem* uma solução e que essa solução é *única*. (Em matemática, essa demonstração é chamada de "demonstração de existência e unicidade".)

Para mostrar que uma solução existe, precisamos apenas verificar que $\mathbf{x} = A^{-1}\mathbf{b}$ funciona. Calculamos

$$A(A^{-1}\mathbf{b}) = (AA^{-1})\mathbf{b} = I\,\mathbf{b} = \mathbf{b}$$

Assim, $A^{-1}\mathbf{b}$ satisfaz a equação $A\mathbf{x} = \mathbf{b}$, que, portanto, tem *pelo menos* essa solução.

Para mostrar que essa solução é única, suponha que $\mathbf{y}$ seja uma outra solução. Então, $A\mathbf{y} = \mathbf{b}$, e, multiplicando à esquerda ambos os lados da equação por A^{-1}, obtemos

$$A^{-1}(A\mathbf{y}) = A^{-1}\mathbf{b} \Rightarrow (A^{-1}A)\mathbf{y} = A^{-1}\mathbf{b} \Rightarrow I\mathbf{y} = A^{-1}\mathbf{b} \Rightarrow \mathbf{y} = A^{-1}\mathbf{b}$$

Portanto, $\mathbf{y}$ é a solução que tínhamos, por isso a solução é única. ◆

Agora, retornando às questões que levantamos nas observações anteriores ao Teorema 1, como podemos saber quando uma matriz tem uma inversa e como podemos achar a inversa quando a matriz é invertível? Daremos brevemente um procedimento geral, mas a situação para matrizes 2×2 é suficientemente simples para ser destacada.

◆ **TEOREMA 3**

Se $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$, então A será invertível se $ad - bc \neq 0$, caso em que

$$A^{-1} = \frac{1}{ad - bc}\begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$$

Se $ad - bc = 0$, A não será invertível.

A expressão $ad - bc$, chamada de ***determinante*** de A, é denotada por det A. A fórmula para a inversa de $\begin{bmatrix} a & b \\ c & d \end{bmatrix}$ (quando ela existe) é, portanto, $\frac{1}{\det A}$ vezes a matriz obtida permutando-se os elementos da diagonal

Material com direitos autorais







You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.

Da mesma forma, $(A^{-1})^2A^2 = I$. Portanto, $(A^{-1})^2$ é a inversa de A^2. Não é difícil ver que um argumento análogo funciona para todos os valores maiores do inteiro n. Entretanto, a indução matemática é a forma de fazer a demonstração corretamente.

O primeiro passo é para $n = 0$, no qual devemos provar que A^0 é invertível e que

$$(A^0)^{-1} = (A^{-1})^0$$

Isso é equivalente a mostrar que I é invertível e que $I^{-1} = I$, o que é claramente verdadeiro.

(Por quê? Veja o Exercício 16.)

Agora assuma que o resultado seja verdadeiro para $n = k$, sendo k um inteiro não negativo específico. A hipótese de indução diz que A^k é invertível e que

$$(A^k)^{-1} = (A^{-1})^k$$

O passo de indução requer que provemos que A^{k+1} é invertível e que $(A^{k+1})^{-1} = (A^{-1})^{k+1}$. Sabemos, por (c), que $A^{k+1} = A^kA$ é invertível, já que A e (por hipótese de indução) A^k são invertíveis. Mais ainda,

$$\begin{aligned} (A^{-1})^{k+1} &= (A^{-1})^kA^{-1} \\ &= (A^k)^{-1}A^{-1} && \text{Por hipótese de indução} \\ &= (AA^k)^{-1} && \text{Pela propriedade (c)} \\ &= (A^{k+1})^{-1} \end{aligned}$$

Portanto, A^n é invertível para todo inteiro não negativo n, e $(A^n)^{-1} = (A^{-1})^n$, pelo princípio da indução matemática. ◆

Observações:

- Apesar de todas as propriedades do Teorema 4 serem úteis, a (c) é a de maior destaque: ela é talvez a propriedade algébrica mais importante das matrizes invertíveis. É também a mais fácil de ser aplicada erroneamente. No Exercício 17, você deve dar um contra-exemplo para mostrar que, ao contrário do que imaginávamos, em geral $(AB)^{-1} \neq A^{-1}B^{-1}$. A propriedade correta, $(AB)^{-1} = B^{-1}A^{-1}$, é algumas vezes chamada de regra da meia e sapato: apesar de calçarmos as meias antes dos sapatos, nós os retiramos na ordem inversa.
- A propriedade (c) é generalizada para produtos de um número finito de matrizes invertíveis: se $A_1, A_2, \ldots, A_n$ são matrizes invertíveis de mesma ordem, então $A_1A_2 \ldots A_n$ é invertível e

$$(A_1A_2 \cdots A_n)^{-1} = A_n^{-1} \cdots A_2^{-1}A_1^{-1}$$

(Veja o Exercício 18.) Assim, podemos enunciar:

A inversa de um produto de matrizes invertíveis é o produto das inversas na ordem reversa.

- Como, para números reais, $\frac{1}{a+b} \neq \frac{1}{a} + \frac{1}{b}$, não devemos esperar que, para matrizes quadradas, $(A + B)^{-1} = A^{-1} + B^{-1}$ (e isso realmente não é verdade em geral; veja o Exercício 19). Na realidade, exceto para matrizes especiais, não existe nenhuma fórmula para $(A + B)^{-1}$.
- A propriedade (e) nos permite definir potências inteiras negativas de uma matriz invertível:

Material com direitos autorais











para o problema de encontrar A^{-1} através da resolução da equação matricial $AX = I_n$ para uma matriz X $n \times n$. (Isso é suficiente pelo Teorema Fundamental, já que uma inversa de A à direita deve ser a sua inversa.) Se denotarmos as colunas de X por $\mathbf{x}_1, \ldots, \mathbf{x}_n$, essa equação matricial será equivalente à resolução nas colunas de X, uma de cada vez. Agora, sendo as colunas de I_n os vetores canônicos unitários $\mathbf{e}_1, \ldots, \mathbf{e}_n$, temos n sistemas de equações lineares, todos com A sendo a matriz dos coeficientes:

$$A\mathbf{x}_1 = \mathbf{e}_1, \ldots, A\mathbf{x}_n = \mathbf{e}_n$$

Como é necessária a mesma seqüência de operações linha para transformar A na sua forma escalonada reduzida, as matrizes completas desses sistemas, $[A \mid \mathbf{e}_1], \ldots, [A \mid \mathbf{e}_n]$, podem ser combinadas por

$$[A \mid \mathbf{e}_1\, \mathbf{e}_2 \cdots \mathbf{e}_n] = [A \mid I_n]$$

Agora efetuamos operações linha para tentar reduzir A a I_n, as quais, se bem-sucedidas, transformarão I_n em A^{-1}.

Ilustramos com três exemplos esse uso do escalonamento de Gauss-Jordan.

EXEMPLO 8 Encontre a inversa de

$$A = \begin{bmatrix} 1 & 2 & -1 \\ 2 & 2 & 4 \\ 1 & 3 & -3 \end{bmatrix}$$

se ela existir.

SOLUÇÃO: O escalonamento de Gauss-Jordan fornece

$$[A \mid I] = \left[\begin{array}{ccc|ccc} 1 & 2 & -1 & 1 & 0 & 0 \\ 2 & 2 & 4 & 0 & 1 & 0 \\ 1 & 3 & -3 & 0 & 0 & 1 \end{array}\right]$$

$$\xrightarrow[L_3 - L_1]{L_2 - 2L_1} \left[\begin{array}{ccc|ccc} 1 & 2 & -1 & 1 & 0 & 0 \\ 0 & -2 & 6 & -2 & 1 & 0 \\ 0 & 1 & -2 & -1 & 0 & 1 \end{array}\right]$$

$$\xrightarrow{(-\frac{1}{2})L_2} \left[\begin{array}{ccc|ccc} 1 & 2 & -1 & 1 & 0 & 0 \\ 0 & 1 & -3 & 1 & -\frac{1}{2} & 0 \\ 0 & 1 & -2 & -1 & 0 & 1 \end{array}\right]$$

$$\xrightarrow{L_3 - L_2} \left[\begin{array}{ccc|ccc} 1 & 2 & -1 & 1 & 0 & 0 \\ 0 & 1 & -3 & 1 & -\frac{1}{2} & 0 \\ 0 & 0 & 1 & -2 & \frac{1}{2} & 1 \end{array}\right]$$

$$\xrightarrow[L_2 + 3L_3]{L_1 + L_3} \left[\begin{array}{ccc|ccc} 1 & 2 & 0 & -1 & \frac{1}{2} & 1 \\ 0 & 1 & 0 & -5 & 1 & 3 \\ 0 & 0 & 1 & -2 & \frac{1}{2} & 1 \end{array}\right]$$

$$\xrightarrow{L_1 - 2L_2} \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 9 & -\frac{3}{2} & -5 \\ 0 & 1 & 0 & -5 & 1 & 3 \\ 0 & 0 & 1 & -2 & \frac{1}{2} & 1 \end{array}\right]$$

Material com direitos autorais

Portanto,

$$A^{-1} = \begin{bmatrix} 9 & -\frac{3}{2} & -5 \\ -5 & 1 & 3 \\ -2 & \frac{1}{2} & 1 \end{bmatrix}$$

(Você deve sempre verificar que $AA^{-1} = I$ pela multiplicação direta. Pelo Teorema 8, não precisamos verificar que $A^{-1}A = I$.) ◆

Observação: Note que usamos uma variante do escalonamento de Gauss-Jordan que introduz primeiro todos os zeros *abaixo* dos pivôs 1, da esquerda para a direita e de cima para baixo, e depois conseguimos os zeros *acima* dos pivôs 1, da direita para a esquerda e de baixo para cima. Essa forma economiza cálculos, como já observamos no Capítulo 2, mas você pode achar mais fácil, quando fizer as contas à mão, criar *todos* os zeros em cada coluna conforme for trabalhando. A resposta obviamente será a mesma.

EXEMPLO 9 Encontre a inversa de

$$A = \begin{bmatrix} 2 & 1 & -4 \\ -4 & -1 & 6 \\ -2 & 2 & -2 \end{bmatrix}$$

se ela existir.

SOLUÇÃO: Procederemos como no Exemplo 8, colocando a matriz identidade ao lado da matriz A e manipulando $[A \mid I]$ tentando transformá-la em $[I \mid A^{-1}]$.

$$[A|I] = \left[\begin{array}{ccc|ccc} 2 & 1 & -4 & 1 & 0 & 0 \\ -4 & -1 & 6 & 0 & 1 & 0 \\ -2 & 2 & -2 & 0 & 0 & 1 \end{array}\right]$$

$$\xrightarrow[L_3+L_1]{L_2+2L_1} \left[\begin{array}{ccc|ccc} 2 & 1 & -4 & 1 & 0 & 0 \\ 0 & 1 & -2 & 2 & 1 & 0 \\ 0 & 3 & -6 & 1 & 0 & 1 \end{array}\right]$$

$$\xrightarrow{L_3-3L_2} \left[\begin{array}{ccc|ccc} 1 & 2 & -1 & 1 & 0 & 0 \\ 0 & 1 & -3 & 2 & 1 & 0 \\ 0 & 0 & 0 & -5 & -3 & 1 \end{array}\right]$$

Nesse ponto, vemos que não é possível reduzir A a I, já que encontramos uma linha de zeros do lado esquerdo da matriz completa. Conseqüentemente, A não é invertível. ◆

Como o próximo exemplo ilustra, tudo funciona da mesma maneira em $\mathbb{Z}_p$, em que p é um inteiro primo.

EXEMPLO 10 Encontre a inversa de

$$A = \begin{bmatrix} 2 & 2 \\ 0 & 0 \end{bmatrix}$$

sobre $\mathbb{Z}_3$, se ela existir.

Material com direitos autorais

















You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.













You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.















































You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.



Vamos chamar a matriz de P e os vetores, de $\mathbf{x}_0 = \begin{bmatrix} 120 \\ 80 \end{bmatrix}$ e $\mathbf{x}_1 = \begin{bmatrix} 100 \\ 100 \end{bmatrix}$. (Note que as componentes de cada vetor são o número de usuários das marcas A e B, respectivamente, após o número de meses indicado pelo índice do vetor.) Dessa forma, temos $\mathbf{x}_1 = P\mathbf{x}_0$.

Estendendo a notação, considere $\mathbf{x}_k$ como o vetor cujas componentes registram a distribuição de usuários de cada pasta de dentes após k meses. Para determinar o número de usuários de cada marca após dois meses, simplesmente aplicamos a mesma equação matricial, começando com $\mathbf{x}_1$ em vez de $\mathbf{x}_0$. Obtemos

$$\mathbf{x}_2 = P\mathbf{x}_1 = \begin{bmatrix} 0{,}70 & 0{,}20 \\ 0{,}30 & 0{,}80 \end{bmatrix}\begin{bmatrix} 100 \\ 100 \end{bmatrix} = \begin{bmatrix} 90 \\ 110 \end{bmatrix}$$

na qual vemos que agora temos 90 usuários da marca A e 110 da marca B. ◆

Os vetores $\mathbf{x}_k$ do Exemplo 1 são chamados de ***vetores de estado*** da cadeia de Markov, e a matriz P é chamada de ***matriz de transição***. Acabamos de ver que uma cadeia de Markov satisfaz a relação

$$\mathbf{x}_{k+1} = P\mathbf{x}_k \quad \text{para} \quad k = 0, 1, 2, \ldots$$

Desse resultado, segue que podemos calcular um vetor de estado arbitrário *iterativamente*, uma vez que conhecemos $\mathbf{x}_0$ e P. Em outras palavras, uma cadeia de Markov é *completamente determinada* por suas probabilidades de transição e por seu estado inicial.

Observações:

- Se, no Exemplo 1, quiséssemos acompanhar não o número *real* de usuários, mas o número *relativo* que utiliza cada marca, poderíamos converter os dados em porcentagens ou frações dividindo por 200 o número total de usuários. Dessa maneira, começaríamos com

$$\mathbf{x}_0 = \begin{bmatrix} \frac{120}{200} \\ \frac{80}{200} \end{bmatrix} = \begin{bmatrix} 0{,}60 \\ 0{,}40 \end{bmatrix}$$

  para indicar o fato de que, inicialmente, a divisão entre as marcas A e B seria de 60% e 40%, respectivamente. Verifique, por cálculo direto, que $P\mathbf{x}_0 = \begin{bmatrix} 0{,}50 \\ 0{,}50 \end{bmatrix}$, que pode ser visto como $\mathbf{x}_1$ (de acordo com a divisão 50%–50% que calculamos agora). Vetores como esses, com componentes não negativas que somam 1, são chamados de ***vetores de probabilidade***.

- Observe como as probabilidades de transição estão arranjadas dentro da matriz de transição P. Podemos pensar nas colunas de P como os estados *presentes*, e nas linhas de P como os estados *seguintes*:

$$\textit{Seguinte}\ \begin{matrix} \text{A} \\ \text{B} \end{matrix}\overset{\textit{Presente}}{\begin{bmatrix} 0{,}70 & 0{,}20 \\ 0{,}30 & 0{,}80 \end{bmatrix}}$$

Material com direitos autorais

A palavra *estocástico* é derivada do adjetivo grego *stokhastikos*, que significa "capaz de aproximar" (ou adivinhar). É aplicada em qualquer coisa governada pelas leis da probabilidade, no sentido de que probabilidade faz previsões sobre a chance de as coisas acontecerem. Na teoria da probabilidade, os "processos estocásticos" são uma generalização das cadeias de Markov.

Note também que as colunas de P são vetores de probabilidade; toda matriz quadrada com essa propriedade é chamada de ***matriz estocástica***.

Podemos perceber de outro modo a natureza determinística das cadeias de Markov. Note que podemos escrever

$$\mathbf{x}_2 = P\mathbf{x}_1 = P(P\mathbf{x}_0) = P^2\mathbf{x}_0$$

e, geralmente,

$$\mathbf{x}_k = P^k\mathbf{x}_0 \quad \text{para } k = 0, 1, 2, \ldots$$

Isso nos leva a examinar as potências de uma matriz de transição. No Exemplo 1, temos

$$P^2 = \begin{bmatrix} 0,70 & 0,20 \\ 0,30 & 0,80 \end{bmatrix}\begin{bmatrix} 0,70 & 0,20 \\ 0,30 & 0,80 \end{bmatrix} = \begin{bmatrix} 0,55 & 0,30 \\ 0,45 & 0,70 \end{bmatrix}$$

O que podemos dizer sobre os elementos dessa matriz? A primeira coisa a observar é que P^2 é outra matriz estocástica, já que a soma dos elementos de suas colunas é igual a 1. (Você provará isso no Exercício 14.) P^2 poderia ser também uma matriz de transição de algum tipo? Considere um de seus elementos — digamos, $(P^2)_{21} = 0,45$. O diagrama de árvore na Figura 2 esclarece de onde esse elemento veio.

Figura 2

Quatro mudanças de estado podem ocorrer em dois meses, e elas correspondem aos quatro galhos (ou caminhos) de comprimento 2 na árvore. Alguém, inicialmente usando a marca A, pode terminar usando a marca B dois meses depois, por dois caminhos diferentes (marcados com * na figura): a pessoa pode continuar usando A depois de um mês e então mudar para B (com probabilidade $0,7(0,3) = 0,21$), ou mudar para B depois de um mês e permanecer com ela (com probabilidade $0,3(0,8) = 0,24$). A soma dessas probabilidades nos dá uma probabilidade geral de 0,45. Observe que esses cálculos são *exatamente* o que fazemos quando calculamos $(P^2)_{21}$.

Em conseqüência, $(P^2)_{21} = 0,45$ representa a probabilidade de passar do estado 1 (marca A) para o estado 2 (marca B) em duas transições. (Note que a ordem dos índices é a inversa da que você suporia.) Generalizando o argumento, pode-se mostrar que

$(P^k)_{ij}$ é a probabilidade de se passar do estado j ao estado i em k transições.

No Exemplo 1, o que acontecerá com a distribuição dos usuários de pasta de dentes em um prazo longo? Vamos trabalhar com vetores de probabilidade como vetores de estado. Dando prosseguimento a nossos cálculos (arredondando para três casas decimais), encontramos

Material com direitos autorais











You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.



































podemos definir $\mathbf{y} = U\mathbf{x}$ para obter dois sistemas lineares, tendo cada um deles uma matriz de coeficientes triangular e, portanto, sendo de resolução direta. Primeiro, encontramos $\mathbf{y}$ por *substituição de frente para trás* em $L\mathbf{y} = \mathbf{b}$ (veja os Exercícios 25 e 26 da Seção 2.2). Em seguida, achamos $\mathbf{x}$, que é o que queremos, usando *substituição de trás para a frente* em $U\mathbf{x} = \mathbf{y}$.

7. Use esse método para resolver os seguintes sistemas $A\mathbf{x} = \mathbf{b}$, em que uma fatoração LU é dada.

(a) $A = \begin{bmatrix} -2 & 1 \\ 2 & 5 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ -1 & 1 \end{bmatrix}\begin{bmatrix} -2 & 1 \\ 0 & 6 \end{bmatrix}, \mathbf{b} = \begin{bmatrix} 5 \\ 1 \end{bmatrix}$

(b) $A = \begin{bmatrix} 4 & -2 \\ 2 & 3 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ \frac{1}{2} & 1 \end{bmatrix}\begin{bmatrix} 4 & -2 \\ 0 & 4 \end{bmatrix}, \mathbf{b} = \begin{bmatrix} 0 \\ 8 \end{bmatrix}$

(c) $A = \begin{bmatrix} 2 & 1 & -2 \\ -2 & 3 & -4 \\ 4 & -3 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 2 & -\frac{5}{4} & 1 \end{bmatrix}\begin{bmatrix} 2 & 1 & -2 \\ 0 & 4 & -6 \\ 0 & 0 & -\frac{7}{2} \end{bmatrix}, \mathbf{b} = \begin{bmatrix} -3 \\ 1 \\ 0 \end{bmatrix}$

(d) $A = \begin{bmatrix} 2 & -4 & 0 \\ 3 & -1 & 4 \\ -1 & 2 & 2 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ \frac{3}{2} & 1 & 0 \\ -\frac{1}{2} & 0 & 1 \end{bmatrix}\begin{bmatrix} 2 & -4 & 0 \\ 0 & 5 & 4 \\ 0 & 0 & 2 \end{bmatrix}, \mathbf{b} = \begin{bmatrix} 2 \\ 0 \\ -5 \end{bmatrix}$

Observação: • Do ponto de vista computacional, a fatoração LU é muito compacta, já que podemos *sobrescrever* os elementos de A a partir dos elementos de L e de U assim que eles são calculados. Por exemplo, em (1), você viu que

$$A = \begin{bmatrix} 2 & 1 & 3 \\ 4 & -1 & 3 \\ -2 & 5 & 5 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ -1 & -2 & 1 \end{bmatrix}\begin{bmatrix} 2 & 1 & 3 \\ 0 & -3 & -3 \\ 0 & 0 & 2 \end{bmatrix} = LU$$

Isso pode ser armazenado como

$$\begin{bmatrix} 2 & 1 & 3 \\ 2 & -3 & -3 \\ -1 & -2 & 2 \end{bmatrix}$$

com os elementos colocados na ordem (1, 1), (1, 2), (1, 3), (2, 1), (3, 1), (2, 2), (2, 3), (3, 2), (3, 3). (Verifique que isso funciona!)

A fatoração LU nos dá ainda uma outra visão do processo do método de eliminação de Gauss. Lembre-se de que a expansão em produto externo do produto de duas matrizes expressa esse produto como uma soma de matrizes, tendo cada uma posto 1. Se aplicarmos essa observação à fatoração $A = LU$, veremos que o que o escalonamento pelo método de Gauss faz é escrever sistematicamente A como a soma de matrizes de posto 1. (Veja o Exercício 56 da Seção 3.5.)

8. Use a fatoração LU do Problema 5 para escrever cada uma das matrizes A como uma soma de matrizes de posto 1.

Investigamos agora o problema de adaptar a fatoração LU para lidarmos em casos nos quais são necessárias permutações das linhas durante o escalonamento pelo método de Gauss. Considere a matriz

Material com direitos autorais

$$A = \begin{bmatrix} 1 & 2 & -1 \\ 3 & 6 & 2 \\ -1 & 1 & 4 \end{bmatrix}$$

Um escalonamento direto por linhas produz

que não é uma matriz triangular superior. Entretanto, podemos facilmente converter essa matriz em uma matriz triangular superior apenas trocando as linhas 2 e 3 de B para obtermos

$$U = \begin{bmatrix} 1 & 2 & -1 \\ 0 & 3 & 3 \\ 0 & 0 & 5 \end{bmatrix}$$

Alternativamente, podemos trocar primeiro as linhas 2 e 3 de A. Para isso, considere P como a matriz elementar

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$$

correspondente à permutação das linhas 2 e 3, e E como o produto das matrizes elementares que reduzem PA a U desse modo. Assim, $EPA = U$ e $PA = E^{-1}U = LU$.

9. Prove que, se P é uma matriz elementar correspondente a uma permutação de linhas, $P^{-1} = P^T$.

Provamos que podemos escrever $A = P^{-1}LU = P^TLU$, o que resolve o caso de uma permutação *simples* de linhas. Em geral, P será o produto $P = P_k \ldots P_2P_1$ de todas as matrizes de permutação de linhas $P_1, P_2, \ldots, P_k$ (em que P_1 é obtida em primeiro lugar, e assim por diante). Essa matriz P é chamada de ***matriz de permutação***. Observe que uma matriz de permutação surge da permutação das linhas de uma matriz identidade em alguma ordem.

10. Prove que existem exatamente $n!$ matrizes de permutação $n \times n$.

11. Escreva as matrizes de permutação a seguir como produtos de matrizes (de permutação de linhas) elementares.

(a) $\begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$ (b) $\begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix}$ (c) $\begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}$ d) $\begin{bmatrix} 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 0 \end{bmatrix}$

12. Prove que, se P é uma matriz de permutação, $P^{-1} = P^T$.

Em geral, uma matriz invertível A $n \times n$ tem uma fatoração P^TLU, em que P é uma matriz de permutação, L é triangular inferior unitária e U é triangular superior. Uma vez determinada P, L e U são únicas.

13. Encontre uma fatoração P^TLU para cada uma das matrizes a seguir.

Material com direitos autorais





Problema 2: Agora, *faça uma iteração* do processo descrito no Problema 1 – ou seja, para uma dada escolha de **x**, renomeie os vértices como antes e em seguida aplique A novamente (depois de novo, e de novo) até que surja um padrão. Como as coordenadas dos vetores se tornarão números bastante elevados, faremos uma *mudança de escala*, dividindo cada vetor pelo valor correspondente à sua *maior coordenada* depois de cada iteração. Assim, se os cálculos nos fornecerem o vetor

$$\begin{bmatrix} 4 \\ 2 \\ 1 \\ 1 \end{bmatrix}$$

nós o substituiremos por

$$\frac{1}{4}\begin{bmatrix} 4 \\ 2 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 0{,}5 \\ 0{,}25 \\ 0{,}25 \end{bmatrix}$$

Observe que esse processo garante que a maior coordenada de cada vetor ficará sempre igual a 1. Faça isso para K_4, depois para K_3 e para K_5. Itere pelo menos dez vezes, dando os resultados com precisão de, no mínimo, até a segunda casa decimal. O que parece acontecer?

Problema 3: Você deve ter notado que, em cada caso, os vetores obtidos no processo de iteração se aproximam de um determinado vetor (um vetor em estado estacionário!). Nomeie os vértices do grafo completo com as coordenadas desse vetor e aplique mais uma vez a matriz de vértices A (sem mudanças de escala). Qual a relação entre as novas coordenadas e as anteriores?

Problema 4: Faça uma conjectura para o caso geral K_n. Qual será o vetor em estado estacionário? O que acontecerá se nomearmos K_n com as coordenadas desse último vetor e aplicarmos a matriz de adjacência A, sem mudanças de escala?

Problema 5: Na Figura 2, mostramos o ***grafo de Petersen***. Repita o procedimento dos Problemas de 1 a 3 com esse grafo.

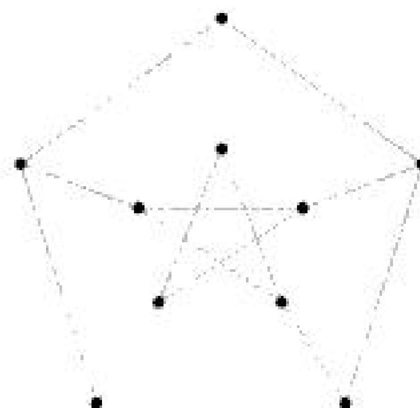


Figura 2

Vamos agora explorar o procedimento descrito anteriormente com outros tipos de grafos para ver se eles se comportam da mesma forma. O ***ciclo*** C_n é o grafo de n vértices apresentado em um formato cíclico. Por exemplo, C_5 é o grafo mostrado na Figura 3.

Problema 6: Repita o procedimento descrito nos Problemas de 1 a 3 com ciclos C_n, para alguns valores ímpares de n, e estabeleça uma conjectura para o caso geral.

Problema 7: Repita o Problema 6, utilizando valores pares de n. O que acontece?

Material com direitos autorais

Um grafo bipartido é um ***grafo bipartido completo*** (veja os Exercícios de 56 a 60 da Seção 3.7) se seus vértices admitem uma partição em conjuntos U e V de maneira que cada vértice de U seja adjacente a todos os vértices de V, e vice-versa. Se U e V tiverem, cada um, n vértices, o grafo será denotado por $K_{n,n}$. Por exemplo, $K_{3,3}$ é o grafo da Figura 4.

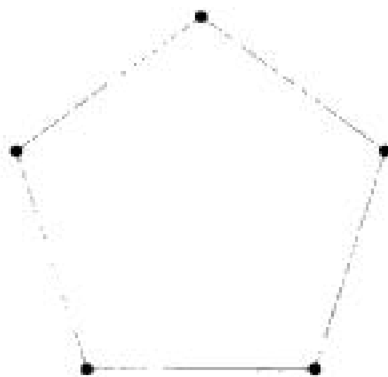


Figura 3

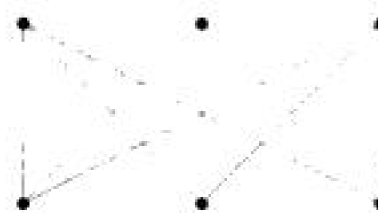


Figura 4

Problema 8: Repita o procedimento descrito nos Problemas de 1 a 3 para grafos bipartidos completos $K_{n,n}$ para vários valores de n. O que acontece?

No final deste capítulo, você terá condições de explicar as observações que fez nesta Introdução.

4.2 Introdução aos Valores e Vetores Próprios

O adjetivo germânico *eigen** significa "próprio" ou "característico de". Valores e vetores próprios, ou autovalores e autovetores, são característicos de uma matriz no sentido de conterem informações importantes sobre a natureza da matriz. A letra λ (lambda), letra grega equivalente ao L em português, é utilizada para designar autovalores porque anteriormente esses números também eram chamados de *valores latentes*. A pronúncia fonética do prefixo alemão *eigen* é "áiguen".**

No Capítulo 3, encontramos a noção de vetor em estado estacionário, no contexto de duas aplicações: cadeias de Markov e modelos de Leslie para o crescimento populacional. Para uma cadeia de Markov com matriz de transição P, um vetor em estado estacionário $\mathbf{x}$ satisfazia a equação $P\mathbf{x} = \mathbf{x}$; para uma matriz de Leslie L, um vetor em estado estacionário correspondia a um vetor população $\mathbf{x}$ satisfazendo $L\mathbf{x} = r\mathbf{x}$, onde r representava a taxa de crescimento do estado estacionário. Por exemplo, vimos que

$$\begin{bmatrix} 0{,}7 & 0{,}2 \\ 0{,}3 & 0{,}8 \end{bmatrix}\begin{bmatrix} 0{,}4 \\ 0{,}6 \end{bmatrix} = \begin{bmatrix} 0{,}4 \\ 0{,}6 \end{bmatrix} \quad \text{e} \quad \begin{bmatrix} 0 & 4 & 3 \\ 0{,}5 & 0 & 0 \\ 0 & 0{,}25 & 0 \end{bmatrix}\begin{bmatrix} 18 \\ 6 \\ 1 \end{bmatrix} = 1{,}5\begin{bmatrix} 18 \\ 6 \\ 1 \end{bmatrix}$$

Neste capítulo, investigamos esse fenômeno de forma mais geral. Para uma matriz quadrada A, investigamos a existência de vetores não nulos $\mathbf{x}$ tais que o valor de $A\mathbf{x}$ seja um múltiplo de $\mathbf{x}$. Esse é o *problema de autovalor*, um dos mais centrais da álgebra linear. Esse tipo de problema tem aplicações por todas as áreas da matemática, e fora dela também.

* N.T.: De ***eigenvalue***, em inglês, e autovalor, em português.

** N.T.: Na literatura especializada em português, também são empregados, como sinônimos de valores e vetores próprios, as expressões valores e vetores característicos, ou autovalores e autovetores. Ao longo do livro, empregaremos preferencialmente esta última designação.

Material com direitos autorais



You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.















Esse método nos dá

$$a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{31}a_{22}a_{13} - a_{32}a_{23}a_{11} - a_{33}a_{21}a_{12}$$

No Exercício 19, você será solicitado a verificar que esse resultado é equivalente ao da equação (1) para um determinante 3×3.

EXEMPLO 2 Calcule o determinante da matriz do Exemplo 1 utilizando o método mostrado em (2).

SOLUÇÃO: Acrescentamos a A suas duas primeiras colunas e calculamos os seis produtos indicados:

$$\begin{array}{ccccc} & & & 0 & -10 \quad -9 \\ \left[\begin{array}{ccc} 5 & -3 & 2 \\ 1 & 0 & 2 \\ 2 & -1 & 3 \end{array}\right] & \begin{array}{cc} 5 & -3 \\ 1 & 0 \\ 2 & -1 \end{array} \\ & & & 0 & -12 \quad -2 \end{array}$$

Somando os três produtos de baixo e subtraindo os três de cima, obtemos

$$\det A = 0 + (-12) + (-2) - 0 - (-10) - (-9) = 5$$ ◆

como antes.

Aviso: Estamos quase definindo determinantes para matrizes arbitrárias. No entanto, *não* existe método análogo ao do Exemplo 2 para matrizes de ordem maior. Ele é válido *somente* para matrizes 3×3.

Determinantes de Matrizes n×n

A definição de determinante de uma matriz 3×3 se estende naturalmente para matrizes quadradas arbitrárias.

Definição Seja $A = [a_{ij}]$ uma matriz $n \times n$, com $n \geq 2$. Então, o ***determinante*** de A é o escalar

$$\det A = |A| = a_{11} \det A_{11} - a_{12} \det A_{12} + \cdots + (-1)^{1+n} a_{1n} \det A_{1n} \qquad (3)$$
$$= \sum_{j=1}^{n} (-1)^{1+j} a_{1j} \det A_{1j}$$

É conveniente combinar um determinante menor complementar com o seu sinal. Com essa finalidade, definimos ***(i, j)-cofator de A*** como

$$C_{ij} = (-1)^{i+j} \det A_{ij}$$

Com essa notação, a definição (3) pode ser reescrita como

$$\det A = \sum_{j=1}^{n} a_{1j} C_{1j} \qquad (4)$$

O Exercício 20 solicitará que você verifique que essa definição fornece corretamente a fórmula para o determinante de uma matriz 2×2 quando $n = 2$.

Material com direitos autorais













You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.















Assim,

$$\det A = (-1)^{i-1} \det B = (-1)^{i-1} \sum_{j=1}^{n} (-1)^{1+j} b_{1j} \det B_{1j}$$
$$= (-1)^{i-1} \sum_{j=1}^{n} (-1)^{1+j} a_{ij} \det A_{ij} = \sum_{j=1}^{n} (-1)^{i+j} a_{ij} \det A_{ij}$$

o que fornece a fórmula para a expansão de co-fatores do det A pela linha i.

A prova para a expansão por coluna é análoga, envolvendo o Lema 13 para que possamos usar a expansão por coluna no lugar da expansão por linha (veja o Exercício 68). ◆

Uma Breve História dos Determinantes

Como observado no início desta seção, a história dos determinantes precede a das matrizes. De fato, determinantes foram introduzidos primeiro, de forma independente, por Seki, em 1683, e por Leibniz, em 1693. Em 1748, determinantes apareceram no *Tratado sobre álgebra*, de Maclaurin, incluindo um tratamento da Regra de Cramer até o caso 4×4. Em 1750, o próprio Cramer provou o caso geral de sua regra, aplicando-a ao ajuste de curvas, e, em 1772, Laplace deu uma prova de seu teorema de expansão.

O termo *determinante* só foi cunhado em 1801, quando utilizado por Gauss. Cauchy foi o primeiro a usar determinantes no sentido moderno, em 1812. Ele foi, de fato, o responsável pelo desenvolvimento de grande parte da teoria inicial de determinantes, incluindo muitos resultados importantes que já mencionamos: a regra do produto para determinantes, o polinômio característico e a noção de matriz diagonalizável. Determinantes só ficaram amplamente conhecidos a partir de 1841, quando Jacobi os popularizou, embora no contexto de funções de várias variáveis, que encontramos em um segundo curso de cálculo. (Sylvester, por volta de 1850, chamou de "jacobianos" esses tipos de determinantes – termo utilizado até hoje.)

No final do século XIX, a teoria dos determinantes havia se desenvolvido a ponto de livros inteiros serem dedicados a ela, incluindo, em 1867, o livro *Uma teoria elementar de determinantes*, de Dodgson, e uma coleção monumental de cinco volumes de Thomas Muir, que apareceu no início do século XX. Mesmo que sua história seja fascinante, determinantes hoje têm um interesse mais teórico do que prático. A Regra de Cramer é um método incorrigivelmente ineficaz na resolução de sistemas de equações lineares, enquanto métodos numéricos substituíram todo o uso que antes era feito de determinantes no cálculo de autovalores. Os determinantes são, no entanto, empregados para propiciar aos estudantes uma compreensão inicial do polinômio característico (como nas Seções 4.2 e 4.4).

Criança prodígio autodidata, Takakazu Seki Kowa (1642–1708) era descendente de uma família de guerreiros samurais. Além de ter descoberto os determinantes, ele escreveu sobre equações diofantinas, quadrados mágicos e números de Bernoulli (antes de Bernoulli), além de, aparentemente, ter feito descobertas em cálculo.

Gottfried Wilhelm von Leibniz (1646–1716) nasceu em Leipzig e estudou Direito, Teologia, Filosofia e Matemática. É provavelmente mais conhecido por ter desenvolvido (como Newton, independentemente) as principais idéias do cálculo diferencial e integral. No entanto, também são impressionantes suas contribuições para outros ramos da matemática. Ele desenvolveu a noção de determinante, conhecia versões da Regra de Cramer e do Teorema da Expansão de Laplace antes que outros tivessem dado os créditos a eles e lançou os fundamentos para a teoria de matrizes, por meio do trabalho que fez sobre as formas quadráticas. Leibniz foi também o primeiro a desenvolver o sistema de numeração binária da aritmética. Acreditava na importância de uma boa notação e, além da notação familiar para derivadas e integrais, introduziu uma espécie de notação com índices subscritos para os coeficientes de um sistema linear, que é, essencialmente, a notação que usamos hoje.

Charles Lutwidge Dodgson (1832–1898) é muito mais conhecido por seu codinome, Lewis Carroll, com o qual assinou os livros *Alice no país das maravilhas* e *Através do espelho*. Ele também escreveu vários livros de matemática e coleções de quebra-cabeças lógicos.







Investigação

Aplicações Geométricas de Determinantes

Esta investigação revelará algumas das aplicações curiosas dos determinantes à geometria. Em particular, veremos que determinantes estão intimamente relacionados com as fórmulas de área e volume, e podem ser utilizados para produzir equações de retas, de planos e de algumas outras curvas. A maioria dessas idéias surgiu quando a teoria dos determinantes estava sendo desenvolvida como um tema independente.

O Produto Vetorial

Relembre que, na Investigação do Capítulo 1, o produto vetorial entre $\mathbf{u} = \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix}$ e $\mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix}$ foi denotado por $\mathbf{u} \times \mathbf{v}$ e definido por

$$\mathbf{u} \times \mathbf{v} = \begin{bmatrix} u_2v_3 - u_3v_2 \\ u_3v_1 - u_1v_3 \\ u_1v_2 - u_2v_1 \end{bmatrix}$$

Se escrevermos esse produto vetorial na forma $(u_2v_3 - u_3v_2)\mathbf{e}_1 - (u_1v_3 - u_3v_1)\mathbf{e}_2 + (u_1v_2 - u_2v_1)\mathbf{e}_3$, onde $\mathbf{e}_1$, $\mathbf{e}_2$ e $\mathbf{e}_3$ são os vetores da base canônica do $\mathbb{R}^3$, veremos que o *formato* dessa expressão pode ser dado por

$$\mathbf{u} \times \mathbf{v} = \det \begin{bmatrix} \mathbf{e}_1 & u_1 & v_1 \\ \mathbf{e}_2 & u_2 & v_2 \\ \mathbf{e}_3 & u_3 & v_3 \end{bmatrix}$$

se expandirmos o determinante pela primeira coluna. (Esse, evidentemente, não é um determinante propriamente dito, já que $\mathbf{e}_1$, $\mathbf{e}_2$ e $\mathbf{e}_3$ são vetores e não escalares; porém, a última igualdade nos fornece uma maneira de lembrar facilmente a fórmula das coordenadas do produto vetorial, que é meio estranha. Ela também nos permite utilizar propriedades dos determinantes para verificar algumas das propriedades do produto vetorial.)

Vamos agora repassar alguns dos exercícios do Capítulo 1.

1. Utilize a versão com determinante do produto vetorial para calcular $\mathbf{u} \times \mathbf{v}$.

(a) $\mathbf{u} = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}, \mathbf{v} = \begin{bmatrix} 3 \\ -1 \\ 2 \end{bmatrix}$ (b) $\mathbf{u} = \begin{bmatrix} 3 \\ -1 \\ 2 \end{bmatrix}, \mathbf{v} = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}$

(c) $\mathbf{u} = \begin{bmatrix} -1 \\ 2 \\ 3 \end{bmatrix}, \mathbf{v} = \begin{bmatrix} 2 \\ -4 \\ -6 \end{bmatrix}$ (d) $\mathbf{u} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \mathbf{v} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$

2. Se $\mathbf{u} = \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix}, \mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix}$ e $\mathbf{w} = \begin{bmatrix} w_1 \\ w_2 \\ w_3 \end{bmatrix}$, mostre que

$$\mathbf{u} \cdot (\mathbf{v} \times \mathbf{w}) = \det \begin{bmatrix} u_1 & v_1 & w_1 \\ u_2 & v_2 & w_2 \\ u_3 & v_3 & w_3 \end{bmatrix}$$

Material com direitos autorais







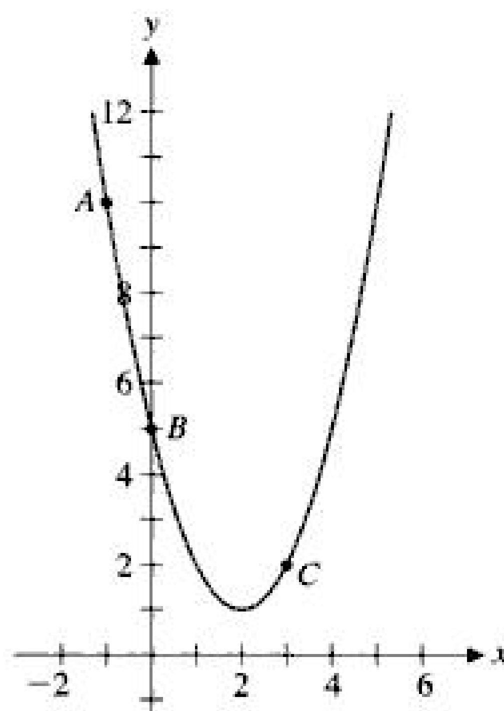


Figura 4

15. A Figura 4 sugere a possibilidade de encontrar uma parábola que passe pelos pontos $A(-1, 10)$, $B(0, 5)$ e $C(3, 2)$. A equação dessa parábola é da forma $y = a + bx + cx^2$. A substituição das coordenadas dos pontos dados nessa equação nos fornece um sistema de três equações lineares nas incógnitas a, b e c. *Sem resolver o sistema*, utilize o Teorema 6 da Seção 4.3 para argumentar que deve haver uma única solução.

Depois, resolva o sistema para encontrar a equação da parábola da Figura 4.

16. Repita o Problema 15 para os pontos $A(-1, -3)$, $B(1, -1)$ e $C(3, 1)$.

17. Generalizando as idéias dos Problemas 15 e 16, suponha que a_1, a_2 e a_3 sejam três números reais distintos. Para quaisquer números reais b_1, b_2 e b_3, queremos mostrar que existe uma única parábola, com equação da forma $y = a + bx + cx^2$, que passa pelos pontos (a_1, b_1), (a_2, b_2) e (a_3, b_3). Faça isso demonstrando que a matriz dos coeficientes do sistema de equações lineares associado tem o determinante

$$\begin{vmatrix} 1 & a_1 & a_1^2 \\ 1 & a_2 & a_2^2 \\ 1 & a_3 & a_3^2 \end{vmatrix} = (a_2 - a_1)(a_3 - a_1)(a_3 - a_2)$$

o qual é necessariamente não nulo. (Por quê?)

18. Sejam a_1, a_2, a_3 e a_4 números reais distintos. Mostre que

$$\begin{vmatrix} 1 & a_1 & a_1^2 & a_1^3 \\ 1 & a_2 & a_2^2 & a_2^3 \\ 1 & a_3 & a_3^2 & a_3^3 \\ 1 & a_4 & a_4^2 & a_4^3 \end{vmatrix} = (a_2 - a_1)(a_3 - a_1)(a_4 - a_1)(a_3 - a_2)(a_4 - a_2)(a_4 - a_3) \neq 0$$

Para quaisquer números reais b_1, b_2, b_3 e b_4, use esse resultado para provar que existe uma única curva cúbica, com equação $y = a + bx + cx^2 + dx^3$ que passa pelos quatro pontos (a_1, b_1), (a_2, b_2), (a_3, b_3) e (a_4, b_4). (*Não* determine os valores de a, b, c e d.)

19. Considere os n números reais $a_1, a_2, ..., a_n$. Prove que

$$\begin{vmatrix} 1 & a_1 & a_1^2 & \cdots & a_1^{n-1} \\ 1 & a_2 & a_2^2 & \cdots & a_2^{n-1} \\ 1 & a_3 & a_3^2 & \cdots & a_3^{n-1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & a_n & a_n^2 & \cdots & a_n^{n-1} \end{vmatrix} = \prod_{1 \le i < j \le n} (a_j - a_i)$$

onde $\Pi_{1 \le i < j \le n}(a_j - a_i)$ representa o produto de todos os termos da forma $(a_j - a_i)$, onde $i < j$ e i e j estão entre 1 e n. [O determinante de uma matriz dessa forma (ou de sua transposta) é chamado de ***determinante de Vandermonde***, nome do matemático francês A. T. Vandermonde (1735–1796).]

Deduza que, para n pontos no plano cartesiano, cujas coordenadas x sejam todas distintas, existe um único polinômio de grau $n - 1$ cujo gráfico passa pelos pontos dados.

Material com direitos autorais

























You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.







You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.



You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.





45. Sejam A e B matrizes semelhantes. Prove que as multiplicidades algébricas dos autovalores de A e de B são iguais.

46. Sejam A e B matrizes semelhantes. Prove que as multiplicidades geométricas dos autovalores de A e de B são iguais. (*Sugestão*: mostre que, se $B = P^{-1}AP$, então todo autovetor de B tem a forma $P^{-1}\mathbf{v}$ para algum autovetor $\mathbf{v}$ de A.)

4.6 Métodos Iterativos para o Cálculo de Autovalores

Até aqui, o único método que temos para o cálculo de autovalores de uma matriz é achar as soluções da sua equação característica. No entanto, esse método enfrenta vários problemas que o tornam praticável somente em um número pequeno de exemplos. O primeiro problema é que ele depende do cálculo de um determinante, o que, para matrizes grandes, é um processo muito demorado. O segundo problema é que a equação característica é uma equação polinomial, e não existem fórmulas para determinar as soluções de equações polinomiais de grau maior do que 4 (equações dos graus 2, 3 e 4 podem ser solucionadas utilizando a fórmula quadrática e suas análogas). Assim, somos forçados a encontrar valores *aproximados* para os autovalores na maioria dos problemas práticos. Infelizmente, métodos para a determinação de raízes aproximadas de uma equação polinomial não são em geral confiáveis por não garantirem uma margem de erro desejável.

Em 1824, o matemático norueguês Niels Henrik Abel (1802–1829) provou que uma equação polinomial genérica do quinto grau (uma qüíntica) não é *solúvel por radicais* – ou seja, não há fórmula que determine suas soluções e que empregue somente as operações de adição, subtração, multiplicação, divisão e extração de raízes *n*-ésimas. Em um artigo escrito em 1830 e publicado postumamente em 1846, o matemático francês Evariste Galois (1811–1832) fez uma teoria mais completa que estabeleceu condições sob as quais uma equação polinomial arbitrária pode ser solúvel por radicais. O trabalho de Galois foi o instrumento que estabeleceu o ramo da álgebra chamado de *teoria de grupos*; sua própria abordagem das equações polinomiais é hoje conhecida como *teoria de Galois*.

Tomaremos um outro caminho para contornar as dificuldades com os polinômios característicos. Vamos primeiro aproximar autovetores e depois usá-los para encontrar seus correspondentes autovalores. Nesta seção, vamos explorar algumas variações de um desses métodos baseado em uma técnica iterativa simples.

O Método da Potência

O método da potência se aplica a matrizes $n \times n$ que tenham um ***autovalor dominante*** λ_1, ou seja, um autovalor cujo valor absoluto seja maior que todos os outros autovalores. Por exemplo, se uma matriz tem autovalores $-4, -3, 1$ e 3, então -4 é o autovalor dominante, já que $4 = |-4| > |-3| \geq |3| \geq |1|$. Por outro lado, uma matriz com autovalores $-4, -3, 3$ e 4 não tem autovalor dominante.

O método da potência procede de forma iterativa para produzir uma seqüência de escalares que converge para λ_1 e uma seqüência de vetores que converge para o autovetor correspondente $\mathbf{v}_1$, o ***autovetor dominante.*** Por simplicidade, vamos assumir que a matriz A é diagonalizável. O Teorema seguinte fornece a base para o método da potência.

◆ **TEOREMA 1**

Seja A uma matriz $n \times n$ diagonalizável com autovalor dominante λ_1. Então, existe um vetor $\mathbf{x}_0$ não nulo tal que a seqüência de vetores $\mathbf{x}_k$ definida por

$$\mathbf{x}_1 = A\mathbf{x}_0, \mathbf{x}_2 = A\mathbf{x}_1, \mathbf{x}_3 = A\mathbf{x}_2, \ldots, \mathbf{x}_k = A\mathbf{x}_{k-1}, \ldots$$

tende a um autovetor dominante de A.

Material com direitos autorais







EXEMPLO 2 Use o método da potência para aproximar o autovalor dominante e um autovetor dominante de

$$A = \begin{bmatrix} 0 & 5 & -6 \\ -4 & 12 & -12 \\ -2 & -2 & 10 \end{bmatrix}$$

SOLUÇÃO: Tomando como vetor inicial

$$\mathbf{x}_0 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

calculamos os dados da Tabela 3.

Tabela 3

k	0	1	2	3	4	5	6	7
$\mathbf{x}_k$	$\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$	$\begin{bmatrix} -1 \\ -4 \\ 6 \end{bmatrix}$	$\begin{bmatrix} -9{,}33 \\ -19{,}33 \\ 11{,}67 \end{bmatrix}$	$\begin{bmatrix} 8{,}62 \\ 17{,}31 \\ -9{,}00 \end{bmatrix}$	$\begin{bmatrix} 8{,}12 \\ 16{,}25 \\ -8{,}20 \end{bmatrix}$	$\begin{bmatrix} 8{,}03 \\ 16{,}05 \\ -8{,}04 \end{bmatrix}$	$\begin{bmatrix} 8{,}01 \\ 16{,}01 \\ -8{,}01 \end{bmatrix}$	$\begin{bmatrix} 8{,}00 \\ 16{,}00 \\ -8{,}00 \end{bmatrix}$
$\mathbf{y}_k$	$\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$	$\begin{bmatrix} -0{,}17 \\ -0{,}67 \\ 1 \end{bmatrix}$	$\begin{bmatrix} 0{,}48 \\ 1 \\ -0{,}60 \end{bmatrix}$	$\begin{bmatrix} 0{,}50 \\ 1 \\ -0{,}52 \end{bmatrix}$	$\begin{bmatrix} 0{,}50 \\ 1 \\ -0{,}50 \end{bmatrix}$	$\begin{bmatrix} 0{,}50 \\ 1 \\ -0{,}50 \end{bmatrix}$	$\begin{bmatrix} 0{,}50 \\ 1 \\ -0{,}50 \end{bmatrix}$	$\begin{bmatrix} 0{,}50 \\ 1 \\ -0{,}50 \end{bmatrix}$
m_k	1	6	−19,33	17,31	16,25	16,05	16,01	16,00

Pode-se ver que os vetores $\mathbf{y}_k$ estão se aproximando de $\begin{bmatrix} 0{,}50 \\ 1 \\ -0{,}50 \end{bmatrix}$, e os escalares m_k, de 16. Isso sugere que eles são, respectivamente, um autovetor dominante e o autovalor dominante de A.

Observações:

- Se o vetor inicial $\mathbf{x}_0$ tiver uma componente nula na direção do autovetor dominante $\mathbf{v}_1$ (isto é, se $c = 0$ na demonstração do Teorema 1), então o método da potência não irá convergir para um autovetor dominante. No entanto, freqüentemente, ao longo das iterações, os cálculos produzirão um $\mathbf{x}_k$ como componente não nula na direção de $\mathbf{v}_1$. A partir daí, o método da potência começará a convergir para um múltiplo de $\mathbf{v}_1$. (Esse é um caso em que os erros de percurso ajudam!)
- O método da potência ainda funciona, sob certas condições, quando se tem autovalor dominante *repetido*, ou mesmo quando a matriz não é diagonalizável. Mais detalhes podem ser encontrados nos mais modernos livros-texto sobre análise numérica. (Veja os Exercícios de 21 a 24.)
- Para algumas matrizes, o método da potência converge rapidamente para um autovetor dominante; para outras, a convergência pode ser bastante vagarosa. Uma análise cuidadosa da prova do Teorema 1 revela o motivo. Como $|\lambda_2/\lambda_1| \geq |\lambda_3/\lambda_1| \geq \ldots \geq |\lambda_n/\lambda_1|$, se $|\lambda_2/\lambda_1|$ for um valor próximo de zero, então $(\lambda_2/\lambda_1)^k, \ldots, (\lambda_n/\lambda_1)^k$ irá se aproximar de zero rapidamente. A equação (2) mostra que $\mathbf{x}_k = A^k\mathbf{x}_0$ também irá se aproximar rapidamente de $\lambda_1^k c_1 \mathbf{v}_1$. Considere o Exemplo 2 como ilustração. Os valores próprios são 16, 4 e 2, por isso $\lambda_2/\lambda_1 = 4/16 = 0{,}25$. Como $0{,}25^7 \approx 0{,}00006$, na sétima iteração precisamos de valores exatos até a quarta casa decimal. Foi exatamente isso que vimos.
- Há uma maneira alternativa para estimar o autovalor dominante λ_1 de uma matriz A, conjuntamente com o método da potência. Primeiro, observe que, se $A\mathbf{x} = \lambda_1\mathbf{x}$, então

Material com direitos autorais







Nos Exercícios de 5 a 8, são dados uma matriz A e um vetor $\mathbf{x}_k$, produzido por iterações, usando o método da potência, como no Exemplo 2. (a) Encontre aproximações do autovalor e do autovetor dominantes de A, calculando os correspondentes m_k e $\mathbf{y}_k$. (b) Verifique que você encontrou aproximações de um autovalor e de um autovetor de A comparando $A\mathbf{y}_k$ com $m_k\mathbf{y}_k$.

5. $A = \begin{bmatrix} 2 & -3 \\ -3 & 10 \end{bmatrix}, \mathbf{x}_5 = \begin{bmatrix} -3{,}667 \\ 11{,}001 \end{bmatrix}$

6. $A = \begin{bmatrix} 5 & 2 \\ 2 & -2 \end{bmatrix}, \mathbf{x}_{10} = \begin{bmatrix} 5{,}530 \\ 1{,}470 \end{bmatrix}$

7. $A = \begin{bmatrix} 4 & 0 & 6 \\ -1 & 3 & 1 \\ 6 & 0 & 4 \end{bmatrix}, \mathbf{x}_8 = \begin{bmatrix} 10{,}000 \\ 0{,}001 \\ 10{,}000 \end{bmatrix}$

8. $A = \begin{bmatrix} 1 & 2 & -2 \\ 1 & 1 & -3 \\ 0 & -1 & 1 \end{bmatrix}, \mathbf{x}_{10} = \begin{bmatrix} 3{,}415 \\ 2{,}914 \\ -1{,}207 \end{bmatrix}$

Nos Exercícios de 9 a 14, utilize o método da potência para encontrar aproximações do autovalor e do autovetor dominantes de A. Use o vetor inicial $\mathbf{x}_0$ fornecido, faça as k iterações solicitadas e dê seus resultados com precisão de três casas decimais.

9. $A = \begin{bmatrix} 14 & 12 \\ 5 & 3 \end{bmatrix}, \mathbf{x}_0 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, k = 5$

10. $A = \begin{bmatrix} -6 & 4 \\ 8 & -2 \end{bmatrix}, \mathbf{x}_0 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, k = 6$

11. $A = \begin{bmatrix} 7 & 2 \\ 2 & 3 \end{bmatrix}, \mathbf{x}_0 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, k = 6$

12. $A = \begin{bmatrix} 3{,}5 & 1{,}5 \\ 1{,}5 & -0{,}5 \end{bmatrix}, \mathbf{x}_0 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, k = 6$

13. $A = \begin{bmatrix} 9 & 4 & 8 \\ 4 & 15 & -4 \\ 8 & -4 & 9 \end{bmatrix}, \mathbf{x}_0 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, k = 5$

14. $A = \begin{bmatrix} 3 & 1 & 0 \\ 1 & 3 & 1 \\ 0 & 1 & 3 \end{bmatrix}, \mathbf{x}_0 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, k = 6$

Nos Exercícios 15 e 16, utilize o método da potência para encontrar aproximações do autovalor e do autovetor dominantes de A, com uma precisão de duas casas decimais. Escolha qualquer vetor inicial que lhe agrade (mas lembre-se da primeira Observação da página 286 e aplique o método até que o dígito da segunda casa decimal, nas iterações, pare de mudar.

15. $A = \begin{bmatrix} 4 & 1 & 3 \\ 0 & 2 & 0 \\ 1 & 1 & 2 \end{bmatrix}$

16. $A = \begin{bmatrix} 12 & 6 & -6 \\ 2 & 0 & -2 \\ -6 & 6 & 12 \end{bmatrix}$

Os quocientes de Rayleigh estão descritos na quarta Observação da página 286. Nos Exercícios de 17 a 20, para ver como o método dos quocientes de Rayleigh aproxima mais rapidamente o autovalor dominante, em comparação ao método da potência ordinário, calcule os sucessivos quocientes de Rayleigh $R(\mathbf{x}_i)$ com $i = 1, \ldots, k$ para a matriz A do exercício mencionado.

17. Exercício 11

18. Exercício 12

19. Exercício 13

20. Exercício 14

As matrizes dos Exercícios de 21 a 24 ou não são diagonalizáveis ou não têm um autovalor dominante (ou as duas coisas). Aplique algum método da potência com o vetor inicial $\mathbf{x}_0$ fornecido, fazendo, em cada caso, oito iterações. Calcule exatamente os autovalores e os autovetores e explique o que está acontecendo.

21. $A = \begin{bmatrix} 4 & 1 \\ 0 & 4 \end{bmatrix}, \mathbf{x}_0 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$

22. $A = \begin{bmatrix} 3 & 1 \\ -1 & 1 \end{bmatrix}, \mathbf{x}_0 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$

23. $A = \begin{bmatrix} 4 & 0 & 1 \\ 0 & 4 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \mathbf{x}_0 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$

24. $A = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 5 & 1 \\ 0 & 0 & 5 \end{bmatrix}, \mathbf{x}_0 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$

Nos Exercícios de 25 a 28, o método da potência não converge para o autovalor e o autovetor dominantes. Verifique isso, utilizando o vetor inicial $\mathbf{x}_0$ dado. Calcule os autovalores e autovetores exatos e dê uma explicação para o que está ocorrendo.

25. $A = \begin{bmatrix} -1 & 2 \\ -1 & 1 \end{bmatrix}, \mathbf{x}_0 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$

26. $A = \begin{bmatrix} 2 & 1 \\ -2 & 5 \end{bmatrix}, \mathbf{x}_0 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$

27. $A = \begin{bmatrix} -5 & 1 & 7 \\ 0 & 4 & 0 \\ 7 & 1 & -5 \end{bmatrix}, \mathbf{x}_0 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$

28. $A = \begin{bmatrix} 1 & -1 & 0 \\ 1 & 1 & 0 \\ 1 & -1 & 1 \end{bmatrix}, \mathbf{x}_0 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$

Material com direitos autorais



































You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.



You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.



























You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.



(b) $x - 2z = 0$ com $\mathbf{u}_1 = \begin{bmatrix} 2/\sqrt{5} \\ 0 \\ 1/\sqrt{5} \end{bmatrix}$ e $\mathbf{u}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$

(c) $2x - 3y + z = 0$ com $\mathbf{u}_1 = \begin{bmatrix} 1/\sqrt{3} \\ -1/\sqrt{3} \\ 1/\sqrt{3} \end{bmatrix}$ e $\mathbf{u}_2 = \begin{bmatrix} 2/\sqrt{6} \\ 1/\sqrt{6} \\ -1/\sqrt{6} \end{bmatrix}$

Problema 10: Mostre que a matriz de projeção dada pela equação (1) satisfaz as propriedades (a) e (b) do Problema 4.

Problema 11: Mostre que a matriz P de uma projeção sobre um plano em $\mathbb{R}^3$ pode ser expressa como

$$P = AA^T$$

onde A é uma matriz 3×2. (*Sugestão*: mostre que a equação (1) é uma expansão em produto externo.)

Problema 12: Mostre que, se P é a matriz de uma projeção sobre um plano em $\mathbb{R}^3$, posto $(P) = 2$.

Neste capítulo, veremos os conceitos de ortogonalidade e projeção ortogonal com grandes detalhes. Veremos que as idéias introduzidas nesta seção podem ser generalizadas e têm muitas aplicações importantes.

5.2 Ortogonalidade em $\mathbb{R}^n$

Nesta seção, generalizaremos a noção de ortogonalidade de dois vetores em $\mathbb{R}^n$ para um conjunto de vetores. Ao fazer isso, veremos que duas propriedades tornam a base canônica $\{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n\}$ de $\mathbb{R}^n$ fácil de ser manipulada: primeiro, dois vetores distintos quaisquer são ortogonais; segundo, cada vetor no conjunto é unitário. Essas duas propriedades nos levam às noções de bases ortogonais e bases ortonormais — conceitos que poderemos usar em uma grande variedade de aplicações.

Conjuntos Ortogonais e Ortonormais de Vetores

Definição Um conjunto de vetores $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\}$ em $\mathbb{R}^n$ é chamado de ***conjunto ortogonal*** quando todos os pares de vetores distintos no conjunto são ortogonais — isto é, quando

$$\mathbf{v}_i \cdot \mathbf{v}_j = 0 \text{ sempre que } i \neq j \text{ para } i, j = 1, 2, \ldots, k$$

A base canônica $\{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n\}$ de $\mathbb{R}^n$ é um conjunto ortogonal assim como todos os seus subconjuntos. Como o primeiro exemplo ilustra, existem muitas outras possibilidades.

EXEMPLO 1 Mostre que $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ é um subconjunto ortogonal de $\mathbb{R}^3$ se

$$\mathbf{v}_1 = \begin{bmatrix} 2 \\ 1 \\ -1 \end{bmatrix}, \quad \mathbf{v}_2 = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}, \quad \mathbf{v}_3 = \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}$$

Material com direitos autorais

SOLUÇÃO: Precisamos mostrar que todo par de vetores desse conjunto é ortogonal. Isso é verdade, pois

$$\mathbf{v}_1 \cdot \mathbf{v}_2 = 2(0) + 1(1) + (-1)(1) = 0$$
$$\mathbf{v}_2 \cdot \mathbf{v}_3 = 0(1) + 1(-1) + (1)(1) = 0$$
$$\mathbf{v}_1 \cdot \mathbf{v}_3 = 2(1) + 1(-1) + (-1)(1) = 0$$

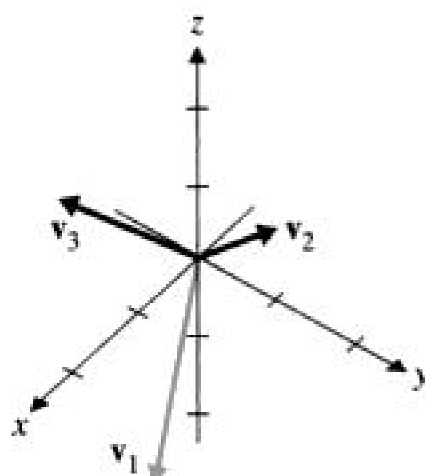


Figura 1 Um conjunto ortogonal de vetores

Geometricamente, os vetores no Exemplo 1 são mutuamente perpendiculares, como mostra a Figura 1. ◆

Uma das vantagens de trabalhar com conjuntos ortogonais de vetores é que eles são linearmente independentes, como mostra o Teorema 1.

◆ TEOREMA 1

Se $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\}$ é um conjunto ortogonal de vetores não nulos em $\mathbb{R}^n$, esses vetores são linearmente independentes.

DEMONSTRAÇÃO: Se $c_1, \ldots, c_k$ são escalares tais que $c_1\mathbf{v}_1 + \ldots + c_k\mathbf{v}_k = \mathbf{0}$, então

$$(c_1\mathbf{v}_1 + \cdots + c_k\mathbf{v}_k) \cdot \mathbf{v}_i = \mathbf{0} \cdot \mathbf{v}_i = 0$$

ou, equivalentemente,

$$c_1(\mathbf{v}_1 \cdot \mathbf{v}_i) + \cdots + c_i(\mathbf{v}_i \cdot \mathbf{v}_i) + \cdots + c_k(\mathbf{v}_k \cdot \mathbf{v}_i) = 0 \qquad (1)$$

Como $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\}$ é um conjunto ortogonal, todos os produtos escalares na equação (1) são iguais a zero, exceto $\mathbf{v}_i \cdot \mathbf{v}_i$. Assim, a equação (1) se reduz a

$$c_i(\mathbf{v}_i \cdot \mathbf{v}_i) = 0$$

Agora, $\mathbf{v}_i \cdot \mathbf{v}_i \neq 0$ pois, por hipótese, $\mathbf{v}_i \neq 0$. Logo, precisamos ter $c_i = 0$. O fato de isso ser verdadeiro para todo $i = 1, \ldots, k$ implica que $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\}$ é um conjunto linearmente independente. ◆

Observação:

- Graças ao Teorema 1, sabemos que, se um conjunto de vetores não nulos é ortogonal, ele é, automaticamente, linearmente independente. Por exemplo, podemos deduzir imediatamente que os três vetores no Exemplo 1 são linearmente independentes. Compare essa abordagem com o trabalho que precisávamos ter para provar a independência linear deles diretamente!

Material com direitos autorais





































Agora, achamos a componente ortogonal de $\mathbf{x}_3$ em relação a

$$W_2 = \text{ger}(\mathbf{x}_1, \mathbf{x}_2) = \text{ger}(\mathbf{v}_1, \mathbf{v}_2) = \text{ger}(\mathbf{v}_1, \mathbf{v}_2')$$

usando a base ortogonal $\{\mathbf{v}_1, \mathbf{v}_2'\}$:

$$\begin{aligned}\mathbf{v}_3 = \text{perp}_{W_2}(\mathbf{x}_3) &= \mathbf{x}_3 - \left(\frac{\mathbf{v}_1 \cdot \mathbf{x}_3}{\mathbf{v}_1 \cdot \mathbf{v}_1}\right)\mathbf{v}_1 - \left(\frac{\mathbf{v}_2' \cdot \mathbf{x}_3}{\mathbf{v}_2' \cdot \mathbf{v}_2'}\right)\mathbf{v}_2' \\ &= \begin{bmatrix} 2 \\ 2 \\ 1 \\ 2 \end{bmatrix} - \left(\tfrac{1}{4}\right)\begin{bmatrix} 1 \\ -1 \\ -1 \\ 1 \end{bmatrix} - \left(\tfrac{15}{20}\right)\begin{bmatrix} 3 \\ 3 \\ 1 \\ 1 \end{bmatrix} \\ &= \begin{bmatrix} -\frac{1}{2} \\ 0 \\ \frac{1}{2} \\ 1 \end{bmatrix}\end{aligned}$$

Multiplicamos agora pelo escalar 2 para obter $\mathbf{v}_3' = 2\mathbf{v}_3 = \begin{bmatrix} -1 \\ 0 \\ 1 \\ 2 \end{bmatrix}$.

Temos agora uma base ortogonal $\{\mathbf{v}_1, \mathbf{v}_2', \mathbf{v}_3'\}$ de W. (Verifique para ter certeza de que esses vetores são ortogonais.) Para obter uma base ortonormal, normalizamos cada vetor:

$$\mathbf{q}_1 = \left(\frac{1}{\|\mathbf{v}_1\|}\right)\mathbf{v}_1 = \left(\frac{1}{2}\right)\begin{bmatrix} 1 \\ -1 \\ -1 \\ 1 \end{bmatrix} = \begin{bmatrix} 1/2 \\ -1/2 \\ -1/2 \\ 1/2 \end{bmatrix}$$

$$\mathbf{q}_2 = \left(\frac{1}{\|\mathbf{v}_2'\|}\right)\mathbf{v}_2' = \left(\frac{1}{2\sqrt{5}}\right)\begin{bmatrix} 3 \\ 3 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 3/2\sqrt{5} \\ 3/2\sqrt{5} \\ 1/2\sqrt{5} \\ 1/2\sqrt{5} \end{bmatrix} = \begin{bmatrix} 3\sqrt{5}/10 \\ 3\sqrt{5}/10 \\ \sqrt{5}/10 \\ \sqrt{5}/10 \end{bmatrix}$$

$$\mathbf{q}_3 = \left(\frac{1}{\|\mathbf{v}_3'\|}\right)\mathbf{v}_3' = \left(\frac{1}{\sqrt{6}}\right)\begin{bmatrix} -1 \\ 0 \\ 1 \\ 2 \end{bmatrix} = \begin{bmatrix} -1/\sqrt{6} \\ 0 \\ 1/\sqrt{6} \\ 2/\sqrt{6} \end{bmatrix} = \begin{bmatrix} -\sqrt{6}/6 \\ 0 \\ \sqrt{6}/6 \\ \sqrt{6}/3 \end{bmatrix}$$

Então, $\{\mathbf{q}_1, \mathbf{q}_2, \mathbf{q}_3\}$ é uma base ortogonal de W. ◆

O Processo de Gram-Schmidt é muito útil para construir uma base ortogonal que contenha um vetor específico. O próximo exemplo ilustra essa aplicação.

EXEMPLO 3 Encontre uma base ortogonal de $\mathbb{R}^3$ que contenha o vetor

$$\mathbf{v}_1 = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$$

Material com direitos autorais







Investigação

Aproximação de Autovalores pelo Algoritmo QR

Um dos melhores (e mais amplamente utilizados) métodos para aproximação numérica de autovalores usa a fatoração QR. O propósito desta investigação é introduzir esse método, o ***algoritmo QR***, e mostrar como ele funciona em alguns exemplos. Para um tratamento mais completo deste tópico, consulte um bom texto sobre álgebra linear numérica.[3] (Você vai achar útil usar um CAS para executar os cálculos nos problemas a seguir.) Dada uma matriz quadrada A, o primeiro passo é fatorar A como $A = QR$ (usando o método que for adequado). Então, definimos $A_1 = RQ$.

1. Primeiro, prove que A_1 é semelhante a A. Depois, prove que A_1 tem os mesmos autovalores de A.

2. Se $A = \begin{bmatrix} 1 & 0 \\ 1 & 3 \end{bmatrix}$, encontre A_1 e verifique que ela tem os mesmos autovalores de A.

Dando continuação ao algoritmo, fatoramos A_1 como $A_1 = Q_1R_1$ e colocamos $A_2 = R_1Q_1$. Então, fatoramos $A_2 = Q_2R_2$ e colocamos $A_3 = R_2Q_2$, e assim por diante. Portanto, para $k \geq 1$, calculamos $A_k = Q_kR_k$ e colocamos $A_{k+1} = R_kQ_k$.

3. Prove que A_k é semelhante a A para todo $k \geq 1$.
4. Dando prosseguimento ao Problema 2, calcule A_2, A_3, A_4 e A_5 com precisão de duas casas decimais. O que você nota?

Pode ser provado que, se os autovalores de A são todos reais e têm valores absolutos distintos, então as matrizes A_k se aproximam de uma matriz triangular superior U.

5. O que valerá para os elementos da diagonal da matriz U?
6. Aproxime os autovalores das matrizes a seguir, aplicando o algoritmo QR. Use precisão de duas casas decimais e execute pelo menos cinco iterações.

(a) $\begin{bmatrix} 2 & 3 \\ 2 & 1 \end{bmatrix}$ (b) $\begin{bmatrix} 1 & 1 \\ 2 & 1 \end{bmatrix}$

(c) $\begin{bmatrix} 1 & 0 & -1 \\ 1 & 2 & 1 \\ -4 & 0 & 1 \end{bmatrix}$ (d) $\begin{bmatrix} 1 & 1 & -1 \\ 0 & 2 & 0 \\ -2 & 4 & 2 \end{bmatrix}$

7. Aplique o algoritmo QR à matriz $A = \begin{bmatrix} 2 & 3 \\ -1 & -2 \end{bmatrix}$. O que acontece? Por quê?

8. Troque os autovalores da matriz do Problema 7 substituindo A por $B = A + 0{,}9I$. Aplique o algoritmo QR à matriz B e então troque de novo, subtraindo 0,9 dos autovalores (aproximados) de B. Verifique que esse método aproxima os autovalores de A.
9. Primeiro mostre que

$$Q_1Q_2 \ldots Q_{k-1}A_k = AQ_1Q_2 \ldots Q_{k-1}$$

[3] Para exemplos, veja G. H. Golub e C. F. Van Loan, *Matrix computations* (Baltimore: Johns Hopkins University Press, 1983).

Material com direitos autorais











































Nos Exercícios de 13 a 16, é dada uma matriz geradora G ou uma matriz de verificação de paridade P para um código C. Encontre uma matriz geradora $G^{\perp}$ e uma matriz de verificação de paridade $P^{\perp}$ para o código dual de C.

13. $G = \begin{bmatrix} 1 & 1 \\ 1 & 1 \\ 1 & 0 \\ 0 & 1 \end{bmatrix}$ **14.** $G = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 0 \\ 1 & 1 \\ 0 & 1 \end{bmatrix}$

15. $P = \begin{bmatrix} 1 & 1 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{bmatrix}$ **16.** $P = \begin{bmatrix} 1 & 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 \end{bmatrix}$

17. Encontre matrizes geradora e de verificação de paridade para o dual do código de Hamming (7, 4) do Exemplo 7 da Seção 3.7.

O ***código de paridade par*** *E_n é o subconjunto de $\mathbb{Z}_2^n$ que consite em todos os vetores de peso par. O* ***código de repetição de n vezes*** *Rep_n é o subconjunto de $\mathbb{Z}_2^n$ que consiste apenas nos dois vetores* **0** *e* **1** *(todas as componentes zero e 1, respectivamente).*

18. (a) Encontre matrizes geradora e de verificação de paridade para E_3 e Rep_3.
(b) Mostre que E_3 e Rep_3 são um o dual do outro.

19. Mostre que E_n e Rep_n são um o dual do outro.

20. Se C e D são códigos com $C \subseteq D$, mostre que $D^{\perp} \subseteq C^{\perp}$.

21. Mostre que, se C é um código com uma matriz geradora, então $(C^{\perp})^{\perp} = C$.

22. Encontre um código autodual de comprimento 6.

Formas Quadráticas

Nos Exercícios de 23 a 28, calcule a forma quadrática $f(\mathbf{x}) = \mathbf{x}^T A\mathbf{x}$ para A e **x** *dados.*

23. $A = \begin{bmatrix} 2 & 3 \\ 3 & 4 \end{bmatrix}, \mathbf{x} = \begin{bmatrix} x \\ y \end{bmatrix}$

24. $A = \begin{bmatrix} 5 & 1 \\ 1 & -1 \end{bmatrix}, \mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$

25. $A = \begin{bmatrix} 3 & -2 \\ -2 & 4 \end{bmatrix}, \mathbf{x} = \begin{bmatrix} 1 \\ 6 \end{bmatrix}$

26. $A = \begin{bmatrix} 1 & 0 & -3 \\ 0 & 2 & 1 \\ -3 & 1 & 3 \end{bmatrix}, \mathbf{x} = \begin{bmatrix} x \\ y \\ z \end{bmatrix}$

27. $A = \begin{bmatrix} 1 & 0 & -3 \\ 0 & 2 & 1 \\ -3 & 1 & 3 \end{bmatrix}, \mathbf{x} = \begin{bmatrix} 2 \\ -1 \\ 1 \end{bmatrix}$

28. $A = \begin{bmatrix} 2 & 2 & 0 \\ 2 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix}, \mathbf{x} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$

Nos Exercícios de 29 a 34, encontre a matriz simétrica A associada à forma quadrática dada.

29. $x_1^2 + 2x_2^2 + 6x_1x_2$ **30.** x_1x_2
31. $3x^2 - 3xy - y^2$ **32.** $x_1^2 - x_3^2 + 8x_1x_2 - 6x_2x_3$
33. $5x_1^2 - x_2^2 + 2x_3^2 + 2x_1x_2 - 4x_1x_3 + 4x_2x_3$
34. $2x^2 - 3y^2 + z^2 - 4xz$

Diagonalize as formas quadráticas nos Exercícios de 35 a 40, encontrando uma matriz ortogonal Q tal que a mudança de variáveis $\mathbf{x} = Q\mathbf{y}$ *transforme a forma dada em uma forma sem produtos mistos. Dê Q e a nova forma quadrática.*

35. $2x_1^2 + 5x_2^2 - 4x_1x_2$ **36.** $x^2 + 8xy + y^2$
37. $7x_1^2 + x_2^2 + x_3^2 + 8x_1x_2 + 8x_1x_3 - 16x_2x_3$
38. $x_1^2 + x_2^2 + 3x_3^2 - 4x_1x_2$ **39.** $x^2 + z^2 - 2xy + 2yz$
40. $2xy + 2xz + 2yz$

Classifique cada uma das formas nos Exercícios de 41 a 48 como positiva definida, positiva semidefinida, negativa definida, negativa semidefinida ou indefinida.

41. $x_1^2 + 2x_2^2$ **42.** $x_1^2 + x_2^2 - 2x_1x_2$
43. $-2x^2 - 2y^2 + 2xy$ **44.** $x^2 + y^2 + 4xy$
45. $2x_1^2 + 2x_2^2 + 2x_3^2 + 2x_1x_2 + 2x_1x_3 + 2x_2x_3$
46. $x_1^2 + x_2^2 + x_3^2 + 2x_1x_3$ **47.** $x_1^2 + x_2^2 - x_3^2 + 4x_1x_2$
48. $-x^2 - y^2 - z^2 - 2xy - 2xz - 2yz$

49. Prove o Teorema 4.

50. Seja $A = \begin{bmatrix} a & b \\ b & d \end{bmatrix}$ uma matriz simétrica 2×2. Prove que A é positiva definida se, e somente se, $a > 0$ e $\det A > 0$. [*Sugestão*:

$$ax^2 + 2bxy + dy^2 = a\left(x + \frac{b}{a}y\right)^2 + \left(d - \frac{b^2}{a}\right)y^2.]$$

Material com direitos autorais









































































$$c_1\begin{bmatrix}1\\1\end{bmatrix} + c_2\begin{bmatrix}2\\3\end{bmatrix} = \begin{bmatrix}-1\\2\end{bmatrix}$$

obtemos $c_1 = -7$ e $c_2 = 3$. Portanto,

$$\begin{aligned} T\begin{bmatrix}-1\\2\end{bmatrix} &= T\left(-7\begin{bmatrix}1\\1\end{bmatrix} + 3\begin{bmatrix}2\\3\end{bmatrix}\right) \\ &= -7T\begin{bmatrix}1\\1\end{bmatrix} + 3T\begin{bmatrix}2\\3\end{bmatrix} \\ &= -7(2 - 3x + x^2) + 3(1 - x^2) \\ &= -11 + 21x - 10x^2 \end{aligned}$$

De maneira similar, descobrimos que

$$\begin{bmatrix}a\\b\end{bmatrix} = (3a - 2b)\begin{bmatrix}1\\1\end{bmatrix} + (b - a)\begin{bmatrix}2\\3\end{bmatrix}$$

portanto,

$$\begin{aligned} T\begin{bmatrix}a\\b\end{bmatrix} &= T\left((3a - 2b)\begin{bmatrix}1\\1\end{bmatrix} + (b - a)\begin{bmatrix}2\\3\end{bmatrix}\right) \\ &= (3a - 2b)T\begin{bmatrix}1\\1\end{bmatrix} + (b - a)T\begin{bmatrix}2\\3\end{bmatrix} \\ &= (3a - 2b)(2 - 3x + x^2) + (b - a)(1 - x^2) \\ &= (5a - 3b) + (-9a + 6b)x + (4a - 3b)x^2 \end{aligned}$$

(Observe que, tomando $a = -1$ e $b = 2$, recuperamos a solução $T\begin{bmatrix}-1\\2\end{bmatrix} = -11 + 21x - 10x^2$.) ◆

A demonstração do teorema geral é imediata.

◆ TEOREMA 2

Seja $T : V \to W$ uma transformação linear, e seja $\mathcal{B} = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ um conjunto gerador de V. Então, $T(\mathcal{B}) = \{T(\mathbf{v}_1), \ldots, T(\mathbf{v}_n)\}$ gera a imagem de T.

DEMONSTRAÇÃO: A imagem de T é o conjunto de todos os vetores de W que são da forma $T(\mathbf{v})$, onde $\mathbf{v}$ é um elemento de V. Seja $T(\mathbf{v})$ na imagem de T. Uma vez que $\mathcal{B}$ gera V, existem escalares $c_1, \ldots, c_n$ tais que

$$\mathbf{v} = c_1\mathbf{v}_1 + \cdots + c_n\mathbf{v}_n$$

Aplicando T e usando o fato de ser uma transformação linear, vemos que

$$T(\mathbf{v}) = T(c_1\mathbf{v}_1 + \cdots + c_n\mathbf{v}_n) = c_1T(\mathbf{v}_1) + \cdots + c_nT(\mathbf{v}_n)$$

Em outras palavras, $T(\mathbf{v})$ pertence ao ger$(T(\mathcal{B}))$, como exigido. ◆

O Teorema 2 se aplica, em particular, quando $\mathcal{B}$ é uma base de V. Você pode estar conjecturando que, neste caso, $T(\mathcal{B})$ seria uma base para a imagem de T. Infelizmente, isso não é sempre verdade. Vamos tratar desse assunto na Seção 6.6.

Material com direitos autorais





























Para isso, sejam $\mathbf{w}_1$ e $\mathbf{w}_2$ em W, e sejam c_1 e c_2 escalares. Como feito anteriormente, sejam $T(\mathbf{v}_1) = \mathbf{w}_1$ e $T(\mathbf{v}_2) = \mathbf{w}_2$. Então $\mathbf{v}_1 = T'(\mathbf{w}_1)$, $\mathbf{v}_2 = T'(\mathbf{w}_2)$ e

$$\begin{aligned} T'(c_1\mathbf{w}_1 + c_2\mathbf{w}_2) &= T'(c_1T(\mathbf{v}_1) + c_2T(\mathbf{v}_2)) \\ &= T'(T(c_1\mathbf{v}_1 + c_2\mathbf{v}_2)) \\ &= I(c_1\mathbf{v}_1 + c_2\mathbf{v}_2) \\ &= c_1\mathbf{v}_1 + c_2\mathbf{v}_2 \\ &= c_1T'(\mathbf{w}_1) + c_2T'(\mathbf{w}_2) \end{aligned}$$

Conseqüentemente, T' é linear e, portanto, pelo Teorema 4 da Seção 6.5, $T' = T^{-1}$. ◆

Isomorfismos de Espaços Vetoriais

Agora estamos em condições de descrever, em termos concretos, o que significa dizer que dois espaços vetoriais são "essencialmente o mesmo".

Definição Uma transformação linear $T : V \to W$ é um ***isomorfismo*** quando é injetora e sobrejetora. Se V e W são dois espaços vetoriais tais que existe um isomorfismo de V em W, então dizemos que V é ***isomorfo*** a W e denotamos por $V \cong W$.

EXEMPLO 14 Prove que $\mathscr{P}_{n-1}$ e $\mathbb{R}^n$ são isomorfos.

SOLUÇÃO: O processo de formar o vetor de coordenadas de um polinômio nos mune de um possível isomorfismo (como já observamos na Seção 6.3, apesar de não termos usado o termo *isomorfismo* naquela ocasião). Especificamente, defina $T : \mathscr{P}_{n-1} \to \mathbb{R}^n$ por $T(p(x)) = [p(x)]_{\mathcal{E}}$, onde $\mathcal{E} = \{1, x, \ldots, x^{n-1}\}$ é a base canônica de $\mathscr{P}_{n-1}$ — isto é,

$$T(a_0 + a_1x + \cdots + a_{n-1}x^{n-1}) = \begin{bmatrix} a_0 \\ a_1 \\ \vdots \\ a_{n-1} \end{bmatrix}$$

O Teorema 3 da Seção 6.3 prova que T é uma transformação linear. Se $p(x) = a_0 + a_1x + \cdots + a_{n-1}x^{n-1}$ está no núcleo de T, então

$$\begin{bmatrix} a_0 \\ \vdots \\ a_{n-1} \end{bmatrix} = T(a_0 + a_1x + \cdots + a_{n-1}x^{n-1}) = \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix}$$

Assim, $a_0 = a_1 = \cdots = a_{n-1} = 0$, e por isso $p(x) = 0$. Portanto, $\ker(T) = \{0\}$, e T é injetora. Como $\dim \mathscr{P}_{n-1} = \dim \mathbb{R}^n = n$, T também é sobrejetora, pelo Teorema 4. Dessa forma, T é um isomorfismo, e $\mathscr{P}_{n-1} \cong \mathbb{R}^n$. ◆

EXEMPLO 15 Prove que M_{mn} e $\mathbb{R}^{mn}$ são isomorfos.

Os termos *isomorfismo* e *isomorfo* são derivados de duas palavras gregas: *isos*, que significa "igual", e *morph*, "forma". Assim, falando figurativamente, espaços vetoriais isomorfos têm "a mesma forma".

SOLUÇÃO: Uma vez mais, a aplicação definida pela formação do vetor de coordenadas de M_{mn} para $\mathbb{R}^{mn}$ (como no Exemplo 15 da Seção 6.3) é um isomorfismo. Os detalhes da demonstração são deixados como exercício. ◆

Material com direitos autorais







Como R é um isomorfismo, ele é invertível; podemos então formar a aplicação composta

$$S \circ T \circ R^{-1} : \mathbb{R}^n \to \mathbb{R}^m$$

que leva $[\mathbf{v}]_{\mathcal{B}}$ em $[T(\mathbf{v})]_{\mathcal{C}}$. Uma vez que essa aplicação vai de $\mathbb{R}^n$ a $\mathbb{R}^m$, sabemos, do Capítulo 3, que é uma transformação matricial. Qual é então a matriz canônica de $S \circ T \circ R^{-1}$? Gostaríamos de encontrar A, a matriz $m \times n$ tal que $A[\mathbf{v}]_{\mathcal{B}} = (S \circ T \circ R^{-1})([\mathbf{v}]_{\mathcal{B}})$. Ou, como $(S \circ T \circ R^{-1})([\mathbf{v}]_{\mathcal{B}}) = [T(\mathbf{v})]_{\mathcal{C}}$, o que queremos é

$$A[\mathbf{v}]_{\mathcal{B}} = [T(\mathbf{v})]_{\mathcal{C}}$$

Surpreendentemente, A é fácil de ser encontrada. A idéia básica é a mesma do Teorema 2 da Seção 3.6. As colunas de A são as imagens dos vetores da base canônica de $\mathbb{R}^n$ pela aplicação de $S \circ T \circ R^{-1}$. Mas, se $\mathcal{B} = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ é uma base para V, então

$$\begin{aligned} R(\mathbf{v}_i) &= [\mathbf{v}_i]_{\mathcal{B}} \\ &= \begin{bmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{bmatrix} \leftarrow i\text{-ésimo elemento} \\ &= \mathbf{e}_i \end{aligned}$$

portanto, $R^{-1}(\mathbf{e}_i) = \mathbf{v}_i$. Conseqüentemente, a i-ésima coluna da matriz A que procuramos é dada por

$$\begin{aligned} (S \circ T \circ R^{-1})(\mathbf{e}_i) &= S(T(R^{-1}(\mathbf{e}_i))) \\ &= S(T(\mathbf{v}_i)) \\ &= [T(\mathbf{v}_i)]_{\mathcal{C}} \end{aligned}$$

que é o vetor de coordenadas de $T(\mathbf{v}_i)$ em relação à base $\mathcal{C}$ de W.

Resumimos essa discussão como um teorema:

◆ TEOREMA 1

Sejam V e W dois espaços vetoriais de dimensão finita com bases $\mathcal{B}$ e $\mathcal{C}$, respectivamente, onde $\mathcal{B} = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$. Se $T : V \to W$ é uma transformação linear, então a matriz A $m \times n$, definida por

$$A = \left[[T(\mathbf{v}_1)]_{\mathcal{C}} \mid [T(\mathbf{v}_2)]_{\mathcal{C}} \mid \cdots \mid [T(\mathbf{v}_n)]_{\mathcal{C}} \right]$$

satisfaz a equação

$$A[\mathbf{v}]_{\mathcal{B}} = [T(\mathbf{v})]_{\mathcal{C}}$$

para todo vetor $\mathbf{v}$ em V.

Material com direitos autorais













You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.



























◆ EXERCÍCIOS 6.8 ◆

Equações Diferenciais Lineares Homogêneas



Nos Exercícios de 1 a 12, encontre a solução da equação diferencial que satisfaz as condições de contorno dadas.

1. $y' - 3y = 0, y(1) = 2$
2. $x' + x = 0, x(1) = 1$
3. $y'' - 7y' + 12y = 0, y(0) = y(1) = 1$
4. $x'' + x' - 12x = 0, x(0) = 0, x'(0) = 1$
5. $f'' - f' - f = 0, f(0) = 0, f(1) = 1$
6. $g'' - 2g = 0, g(0) = 1, g(1) = 0$
7. $y'' - 2y' + y = 0, y(0) = y(1) = 1$
8. $x'' + 4x' + 4x = 0, x(0) = 1, x'(0) = 1$
9. $y'' - k^2y = 0, k \neq 0, y(0) = y'(0) = 1$
10. $y'' - 2ky' + k^2y = 0, k \neq 0, y(0) = 1, y(1) = 0$
11. $f'' - 2f' + 5f = 0, f(0) = 1, f(\pi/4) = 0$
12. $h'' - 4h' + 5h = 0, h(0) = 0, h'(0) = -1$

13. Uma cultura de bactérias tem uma taxa de crescimento proporcional ao tamanho da população. Inicialmente, existem 100 bactérias; depois de três horas, existem 1.600.

(a) Se $p(t)$ denota o número de bactérias após t horas, encontre uma fórmula para $p(t)$.

(b) Quanto tempo demora para a população dobrar de tamanho?

(c) Quando a população atingirá um milhão de indivíduos?

14. A Tabela 1 relaciona a população dos Estados Unidos CAS para os anos entre 1900 e 2000, em intervalos de dez anos.

(a) Assumindo o modelo de crescimento exponencial, use os dados de 1900 e de 1910 para encontrar uma fórmula para $p(t)$, a população no ano t. (*Sugestão:* seja $t = 0$ o ano de 1900, e seja $t = 1$ o ano de 1910.) Quão acuradamente sua fórmula calcula a população norte-americana no ano de 2000?

(b) Repita o item (a), mas use os dados de 1970 e 1980 para encontrar uma fórmula para $p(t)$. Esse enfoque dá uma aproximação melhor para o ano 2000?

(c) O que você pode concluir sobre o crescimento populacional dos Estados Unidos?

15. A meia-vida do rádio-226 é de 1.590 anos. Suponha que comecemos com uma amostra de rádio-226 com massa de 50 mg.

(a) Encontre uma fórmula para a massa $m(t)$ que resta após t anos e use essa fórmula para estimar a quantidade de massa restante após 1.000 anos.

Tabela 1

Ano	População (em milhões)
1900	76
1910	92
1920	106
1930	123
1940	131
1950	150
1960	179
1970	203
1980	227
1990	250
2000	281

Fonte: U.S. Bureau of the Census.

(b) Quando restarão apenas 10 mg?

16. O processo de *datação por carbono radioativo* é usado pelos cientistas para estimar a idade de objetos antigos feitos de matéria viva, como ossos, couro, madeira ou papel. Todos eles contêm carbono, uma parte do qual é carbono-14, um isótopo radioativo continuamente formado nas camadas mais altas da atmosfera. Como os organismos vivos são constituídos de carbono radioativo e de outros átomos de carbono, a proporção entre as duas formas permanece constante. Entretanto, quando um organismo morre, o carbono-14 existente em suas células decai e não é substituído. Sabemos que a meia-vida do carbono-14 é de 5.730 anos, assim, ao medir a concentração de carbono-14 em um objeto, os cientistas podem determinar sua idade aproximada.

Uma das aplicações mais bem-sucedidas da datação por carbono radioativo foi determinar a idade do monumento de Stonehenge, na Inglaterra (Figura 8). Descobriu-se que a concentração de carbono-14, em amostras tiradas dos restos de postes de madeira, era de 45% da que se encontrava em matéria viva. Qual é a idade aproximada desses postes?

Material com direitos autorais























































onde $\varepsilon_i = y_i - (a + bx_i)$. A reta $y = a + bx$ que minimiza $\varepsilon_1^2 + \cdots + \varepsilon_n^2$ é chamada de ***reta de mínimos quadrados*** (ou ***reta de melhor ajuste***) para os pontos $(x_1, y_1), \ldots, (x_n, y_n)$. Como visto antes do Exemplo 2, podemos expressar esse problema em forma de matriz. Se os pontos dados estivessem de fato na reta $y = a + bx$, as n equações lineares

$$\begin{aligned} a + bx_1 &= y_1 \\ &\vdots \\ a + bx_n &= y_n \end{aligned}$$

seriam todas verdadeiras (isto é, o sistema seria compatível). Nosso interesse está no caso em que os pontos *não* são colineares, no qual o sistema é *indeterminado*. Em forma matricial, temos

$$\begin{bmatrix} 1 & x_1 \\ 1 & x_2 \\ \vdots & \vdots \\ 1 & x_n \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}$$

que está na forma $A\mathbf{x} = \mathbf{b}$, onde

$$A = \begin{bmatrix} 1 & x_1 \\ 1 & x_2 \\ \vdots & \vdots \\ 1 & x_n \end{bmatrix}, \quad \mathbf{x} = \begin{bmatrix} a \\ b \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}$$

O vetor-erro $\mathbf{e}$ é apenas $\mathbf{b} - A\mathbf{x}$ (verifique isso), e queremos minimizar $\|\mathbf{e}\|^2$ ou, equivalentemente, $\|\mathbf{e}\|$. Podemos, assim, reescrever nosso problema em forma de matrizes do seguinte modo:

Definição Se A é uma matriz $m \times n$ e $\mathbf{b}$ está em $\mathbb{R}^m$, uma ***solução por mínimos quadrados*** de $A\mathbf{x} = \mathbf{b}$ é um vetor $\bar{\mathbf{x}}$ em $\mathbb{R}^n$ tal que

$$\|\mathbf{b} - A\bar{\mathbf{x}}\| \leq \|\mathbf{b} - A\mathbf{x}\|$$

para todo $\mathbf{x}$ em $\mathbb{R}^n$.

Solução para o Problema dos Mínimos Quadrados

Qualquer vetor da forma $A\mathbf{x}$ está no espaço-coluna de A, e, quando $\mathbf{x}$ varia sobre todos os vetores de $\mathbb{R}^n$, $A\mathbf{x}$ varia sobre todos os vetores em col(A). Uma solução por mínimos quadrados de $A\mathbf{x} = \mathbf{b}$ é, portanto, equivalente a um vetor $\bar{\mathbf{y}}$ em col(A) tal que

$$\|\mathbf{b} - \bar{\mathbf{y}}\| \leq \|\mathbf{b} - \mathbf{y}\|$$

para todo $\mathbf{y}$ em col(A). Em outras palavras, precisamos do vetor em col(A) mais próximo a $\mathbf{b}$. Pelo Teorema da Melhor Aproximação, o vetor que queremos é a projeção ortogonal de $\mathbf{b}$ em col(A). Assim, se $\bar{\mathbf{x}}$ é uma solução por mínimos quadrados de $A\mathbf{x} = \mathbf{b}$, temos

$$A\bar{\mathbf{x}} = \text{proj}_{\text{col}(A)}(\mathbf{b}) \qquad (1)$$

Para acharmos $\bar{\mathbf{x}}$, pode parecer que precisamos primeiro calcular $\text{proj}_{\text{col}(A)}(\mathbf{b})$ e então resolver o sistema (1). Entretanto, há um jeito melhor.

Material com direitos autorais







como mostra a Figura 7.

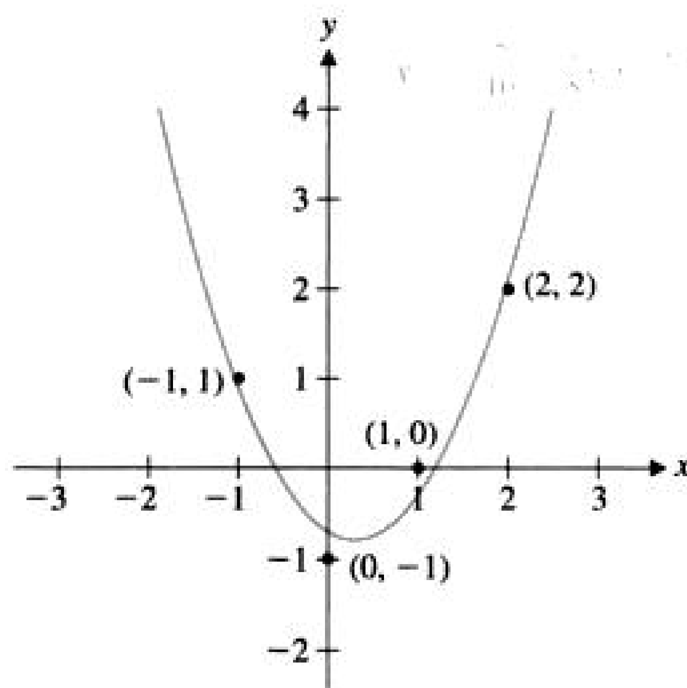


Figura 7 Uma parábola de mínimos quadrados

Uma das importantes aplicações do método dos mínimos quadrados é estimar constantes associadas a vários processos. O próximo exemplo ilustra essa aplicação em um contexto de crescimento populacional. Vimos, na Seção 6.8, que uma população que está crescendo (ou decrescendo) exponencialmente satisfaz uma equação da forma $p(t) = ce^{kt}$, onde $p(t)$ é o tamanho da população no tempo t, e c e k são constantes. Claramente, $c = p(0)$, mas k não é tão fácil de ser determinada. É fácil ver que

$$k = \frac{p'(t)}{p(t)}$$

o que explica por que k é, às vezes, chamada de *taxa de crescimento relativo* da população: ela é a razão da taxa de crescimento $p'(t)$ pelo tamanho da população $p(t)$.

Tabela 2

Ano	População (em bilhões)
1950	2,56
1960	3,04
1970	3,71
1980	4,46
1990	5,28
2000	6,08

Fonte: U. S. Bureau of the Census, International Data Base

CAS

EXEMPLO 7

A Tabela 2 apresenta a população do mundo para intervalos de dez anos, referente à segunda metade do século XX. Supondo um modelo de crescimento exponencial, encontre a taxa de crescimento relativo e preveja a população do mundo para 2010.

Material com direitos autorais





















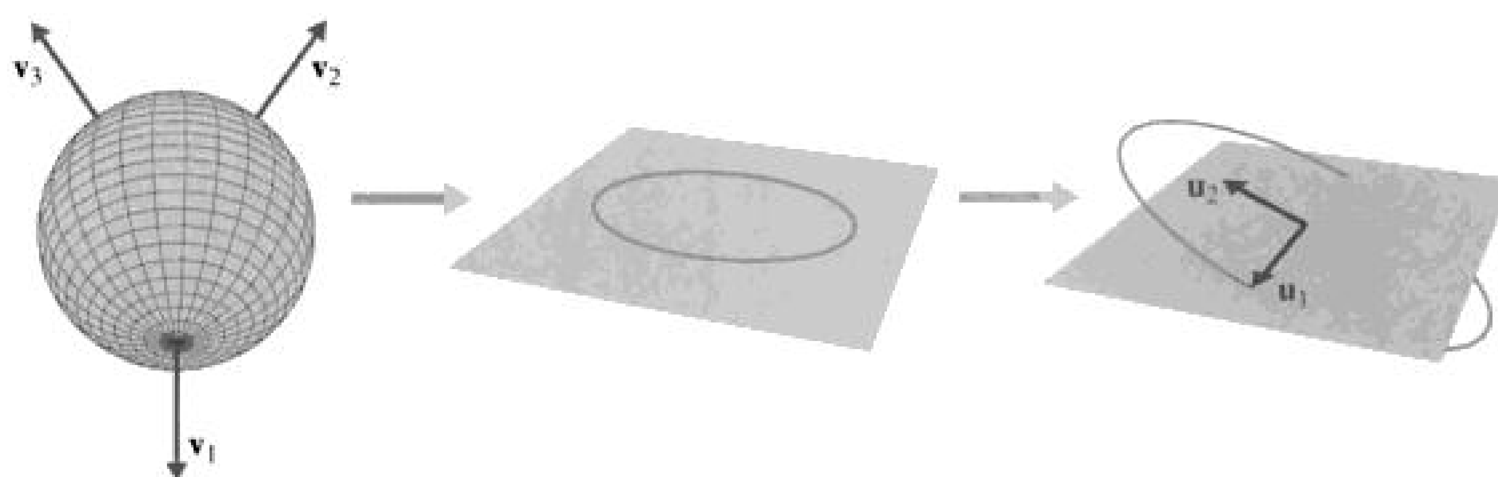


Figura 4

Aplicações da DVS

A decomposição por valores singulares é uma ferramenta extremamente importante, tanto prática quanto teoricamente. Vamos ver apenas algumas das muitas aplicações.

Posto

Até agora, não nos preocupamos em calcular o posto de uma matriz sob um ponto de vista computacional. Calculamos o posto de uma matriz por redução nas linhas para a forma escalonada e contando o número de linhas não nulas. Entretanto, como temos visto, erros de arredondamento podem afetar esse processo, especialmente se a matriz for malcondicionada. Elementos que deveriam ser zero podem acabar sendo números muito pequenos, interferindo em nossa capacidade de determinar com precisão o posto e outras quantidades associadas com a matriz. Na prática, a DVS é freqüentemente utilizada para encontrar o posto de uma matriz, já que ela é muito mais confiável quando erros por arredondamento estão presentes. A idéia básica por trás dessa abordagem é que as matrizes ortogonais U e V da DVS preservam o comprimento e, portanto, não introduzem nenhum erro adicional; qualquer erro que ocorrer tenderá a aparecer na matriz Σ.

CAS **EXEMPLO 4** Sejam

$$A = \begin{bmatrix} 8{,}1650 & -0{,}0041 & -0{,}0041 \\ 4{,}0825 & -3{,}9960 & 4{,}0042 \\ 4{,}0825 & 4{,}0042 & -3{,}9960 \end{bmatrix} \quad \text{e} \quad B = \begin{bmatrix} 8{,}17 & 0 & 0 \\ 4{,}08 & -4 & 4 \\ 4{,}08 & 4 & -4 \end{bmatrix}$$

A matriz B foi obtida ao arredondarmos os valores dos elementos de A para a segunda casa decimal. Se calcularmos os postos dessas duas matrizes aparentemente iguais, encontraremos posto$(A) = 3$, mas posto$(B) = 2$. Pelo Teorema Fundamental, isso implica, entre muitas coisas, que A é invertível, mas B não.

A explicação para essas diferenças críticas entre duas matrizes aproximadamente iguais recaem sobre suas DVSs. Os valores singulares de A são 10, 8 e 0,01, e então A tem posto 3. Os valores singulares de B são 10, 8 e 0, portanto, B tem posto 2.

Em aplicações práticas, é freqüente assumirmos que, se um valor singular calculado for perto de zero, então erros por arredondamento estão aparecendo e o valor real deveria ser zero. Nesse sentido, "ruídos" podem ser filtrados. Neste exemplo, se calcularmos $A = U\Sigma V^T$ e substituirmos

$$\Sigma = \begin{bmatrix} 10 & 0 & 0 \\ 0 & 8 & 0 \\ 0 & 0 & 0{,}01 \end{bmatrix} \quad \text{por} \quad \Sigma' = \begin{bmatrix} 10 & 0 & 0 \\ 0 & 8 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

então, $U\Sigma' V^T = B$. (Tente isso!) ◆

Material com direitos autorais















Jean-Baptiste Joseph Fourier (1768–1830) foi um matemático e físico francês que ganhou reconhecimento por suas pesquisas sobre teoria de calor. Em sua solução limite para a chamada equação de calor, ele introduziu técnicas para o que agora conhecemos como série de Fourier, uma ferramenta bastante utilizada em muitos ramos da matemática, da física e da engenharia. Fourier foi um ativista político durante a Revolução Francesa e se tornou um dos favoritos de Napoleão, acompanhando-o em sua campanha no Egito em 1798. Mais tarde, Napoleão nomeou Fourier prefeito de Isère, onde ele supervisionou muitos projetos importantes de engenharia. Em 1808, Fourier se tornou um barão. Ele é homenageado por uma placa na Torre Eiffel.

Para acharmos a projeção ortogonal de uma função f em $\mathscr{C}[-\pi, \pi]$ sobre o subespaço W gerado pela base ortogonal $\mathcal{B}$, precisamos saber os valores das raízes das normas dos vetores da base. Por exemplo, pela fórmula do ângulo-metade, temos

$$\begin{aligned}\langle \operatorname{sen} kx, \operatorname{sen} kx\rangle &= \int_{-\pi}^{\pi} \operatorname{sen}^2 kx\, dx \\ &= \tfrac{1}{2}\int_{-\pi}^{\pi} (1 - \cos 2kx)\, dx \\ &= \tfrac{1}{2}\left[x - \frac{\operatorname{sen} 2kx}{2k}\right]_{-\pi}^{\pi} \\ &= \pi\end{aligned}$$

No Exercício 20, pediremos que você mostre que $\langle \cos kx, \cos kx\rangle = \pi$ e $\langle 1, 1\rangle = 2\pi$.

Agora, temos

$$\operatorname{proj}_W(f) = a_0 + a_1 \cos x + \cdots + a_n \cos nx + b_1 \operatorname{sen} x + \cdots + b_n \operatorname{sen} nx \qquad (2)$$

onde

$$\begin{aligned}a_0 &= \frac{\langle 1, f\rangle}{\langle 1, 1\rangle} = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x)\, dx \\ a_k &= \frac{\langle \cos kx, f\rangle}{\langle \cos kx, \cos kx\rangle} = \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)\cos kx\, dx \\ b_k &= \frac{\langle \operatorname{sen} kx, f\rangle}{\langle \operatorname{sen} kx, \operatorname{sen} kx\rangle} = \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)\operatorname{sen} kx\, dx\end{aligned} \qquad (3)$$

para $k \geq 1$. A aproximação para f dada pelas equações (2) e (3) é chamada de ***aproximação de Fourier de ordem n*** para f em $[-\pi, \pi]$. Os coeficientes $a_0, a_1, \ldots, a_n, b_1, \ldots, b_n$ são chamados de ***coeficientes de Fourier*** para f.

EXEMPLO 4 Encontre a aproximação de Fourier de ordem 4 para $f(x) = x$ em $[-\pi, \pi]$.

SOLUÇÃO: Pelas fórmulas em (3), temos

$$a_0 = \frac{1}{2\pi}\int_{-\pi}^{\pi} x\, dx = \frac{1}{2\pi}\left[\frac{x^2}{2}\right]_{-\pi}^{\pi} = 0$$

e, para $k \geq 1$, integrando por partes, obtemos

$$a_k = \frac{1}{\pi}\int_{-\pi}^{\pi} x\cos kx\, dx = \frac{1}{\pi}\left[\frac{x}{k}\operatorname{sen} kx + \frac{1}{k^2}\cos kx\right]_{-\pi}^{\pi} = 0$$

e

$$\begin{aligned}b_k &= \frac{1}{\pi}\int_{-\pi}^{\pi} x\operatorname{sen} kx\, dx = \frac{1}{\pi}\left[-\frac{x}{k}\cos kx + \frac{1}{k^2}\operatorname{sen} kx\right]_{-\pi}^{\pi} \\ &= \frac{1}{\pi}\left[\frac{-\pi\cos k\pi - \pi\cos(-k\pi)}{k}\right]\end{aligned}$$

Material com direitos autorais







O conjunto de todos os inteiros é denotado por $\mathbb{Z}$, portanto,

$$\mathbb{Z} = \{\ldots, -2, -1, 0, 1, 2, \ldots\}$$

Dois conjuntos são considerados ***iguais*** se eles contêm exatamente os mesmos elementos. A *ordem* na qual os elementos são listados não importa, e repetições não são contadas. Portanto,

$$\{1, 2, 3\} = \{2, 1, 3\} = \{1, 3, 2, 1\}$$

O símbolo $\in$ significa "é um elemento de" ou "pertence a", e o símbolo $\notin$ denota a negação — ou seja, "não é um elemento de" ou "não pertence a". Por exemplo,

$$5 \in \mathbb{Z}^+ \text{ mas } 0 \notin \mathbb{Z}^+$$

Muitas vezes é mais conveniente descrever um conjunto em termos de uma regra satisfeita por todos os seus elementos. Em tais casos, a ***notação de construção de conjunto*** é apropriada. O formato é

$$\{x : x \text{ satisfaz } P\}$$

onde P representa uma propriedade ou uma coleção de propriedades que o elemento x deve satisfazer. Os dois pontos são pronunciados "tal que". Por exemplo,

$$\{n : n \in \mathbb{Z}, n > 0\}$$

é lido como "o conjunto de todo n tal que n é um inteiro e n é maior que zero". Esse é só mais um meio de descrever os inteiros positivos $\mathbb{Z}^+$. (Também poderíamos escrever $\mathbb{Z}^+ = \{n \in \mathbb{Z} : n > 0\}$.)

O ***conjunto vazio*** é o conjunto sem nenhum elemento. Ele é denotado por $\varnothing$ ou $\{\ \}$.

EXEMPLO 1 Descreva em palavras os seguintes conjuntos:

(a) $A = \{n : n = 2k, k \in \mathbb{Z}\}$ (b) $B = \{m/n : m, n \in \mathbb{Z}, n \neq 0\}$
(c) $C = \{x \in \mathbb{R} : 4x^2 - 4x - 3 = 0\}$ (d) $D = \{x \in \mathbb{Z} : 4x^2 - 4x - 3 = 0\}$

SOLUÇÃO: (a) A é o conjunto dos números n que são múltiplos inteiros de 2. Portanto, A é o conjunto de todos os números inteiros pares.
(b) B é o conjunto de todas as expressões da forma m/n, onde m e n são inteiros e n é não nulo. Esse é o conjunto dos *números racionais*, geralmente denotado por $\mathbb{Q}$. (Note que essa forma de descrever $\mathbb{Q}$ produz muitas repetições; entretanto, nossa convenção, como observamos antes, é a de que só incluímos uma ocorrência para cada elemento. Assim, a expressão descreve precisamente o conjunto de todos os números racionais.)
(c) C é o conjunto de todas as soluções reais da equação $4x^2 - 4x - 3 = 0$. Fatorando ou usando a fórmula quadrática, vemos que as raízes dessa equação são $-\frac{1}{2}$ e $\frac{3}{2}$. (Verifique isso.) Portanto,

$$C = \{-\tfrac{1}{2}, \tfrac{3}{2}\}$$

(d) A partir da solução de (c), vemos que *não* existem soluções para $4x^2 - 4x - 3 = 0$ em $\mathbb{R}$ que sejam inteiras. Portanto, D é o conjunto vazio, o qual podemos expressar por $D = \varnothing$.

Material com direitos autorais







(c) O padrão aqui é menos óbvio, mas um pouco de reflexão revela que cada termo é um a menos que um quadrado perfeito:

$$\begin{aligned} 3 + 8 + 15 + \cdots + 99 &= (2^2 - 1) + (3^2 - 1) + (4^2 - 1) + \cdots + (10^2 - 1) \\ &= \sum_{k=2}^{10} (k^2 - 1) \end{aligned}$$

◆

EXEMPLO 6 Reescreva cada uma das somas do Exemplo 5 de modo que o índice do somatório comece em 1.

SOLUÇÃO:

(a) Se fizermos a mudança de variável $i = k + 1$, então, conforme k variar de 0 a 6, i variará de 1 a 7. Como $k = i - 1$, obteremos

$$\sum_{k=0}^{6} 2^k = \sum_{i=1}^{7} 2^{i-1}$$

(b) Utilizando a mesma substituição do item (a), obtemos

$$\sum_{k=0}^{49} (2k + 1) = \sum_{i=1}^{50} (2(i - 1) + 1) = \sum_{i=1}^{50} (2i - 1)$$

(c) A substituição $i = k - 2$ funcionará (experimente), mas é mais fácil só acrescentar um termo correspondente a $k = 1$, já que $1^2 - 1 = 0$. Portanto,

$$\sum_{k=2}^{10} (k^2 - 1) = \sum_{k=1}^{10} (k^2 - 1)$$

◆

Somas múltiplas aparecem quando existe mais de um índice de somatório, como acontece com uma matriz. A notação

$$\sum_{i,j=1}^{n} a_{ij} \tag{3}$$

significa que é para somar os termos a_{ij} conforme i e j variam independentemente de 1 a n. A soma na expressão (3) é equivalente tanto a

$$\sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij}$$

onde somamos primeiro em j e depois em i (sempre trabalhamos de dentro para fora), quanto a

$$\sum_{j=1}^{n} \sum_{i=1}^{n} a_{ij}$$

onde a ordem da soma é invertida.

EXEMPLO 7 Escreva $\sum_{i,j=1}^{3} i^j$ usando as duas possíveis ordens de soma.

SOLUÇÃO:

$$\begin{aligned} \sum_{i=1}^{3} \sum_{j=1}^{3} i^j &= \sum_{i=1}^{n} (i^1 + i^2 + i^3) \\ &= (1^1 + 1^2 + 1^3) + (2^1 + 2^2 + 2^3) + (3^1 + 3^2 + 3^3) \end{aligned}$$

Material com direitos autorais







Isso completa o passo de indução, e concluímos, pelo Princípio de Indução Matemática, que a fórmula (1) é verdadeira para todo $n \geq 1$. ◆

O próximo exemplo demonstra uma fórmula útil para a soma dos n primeiros inteiros positivos. A fórmula aparece com freqüência no texto; por exemplo, veja a solução do Exercício 45 na Seção 2.5.

EXEMPLO 2 Prove que

$$1 + 2 + \cdots + n = \frac{n(n+1)}{2}$$

para todo $n \geq 1$.

SOLUÇÃO: A fórmula é verdadeira para $n = 1$, pois

$$1 = \frac{1(1+1)}{2}$$

Suponha que a fórmula seja verdadeira para $n = k$, isto é,

$$1 + 2 + \cdots + k = \frac{k(k+1)}{2}$$

Precisamos mostrar que a fórmula vale quando $n = k + 1$, ou seja, precisamos provar que

$$1 + 2 + \cdots + (k+1) = \frac{(k+1)[(k+1)+1]}{2}$$

Mas observe que

$$\begin{aligned} 1 + 2 + \cdots + (k+1) &= (1 + 2 + \cdots + k) + (k+1) \\ &= \frac{k(k+1)}{2} + (k+1) \qquad \text{Pela hipótese de indução} \\ &= \frac{k(k+1) + 2(k+1)}{2} \\ &= \frac{k^2 + 3k + 2}{2} \\ &= \frac{(k+1)(k+2)}{2} \\ &= \frac{(k+1)[(k+1)+1]}{2} \end{aligned}$$

que é o que precisávamos provar.

Isso completa o passo de indução, e concluímos, pelo Princípio de Indução Matemática, que a fórmula é verdadeira para todo $n \geq 1$. ◆

Por um processo análogo, podemos provar que a soma dos quadrados dos n primeiros inteiros positivos satisfaz a fórmula

$$1^2 + 2^2 + 3^2 + \cdots + n^2 = \frac{n(n+1)(2n+1)}{6}$$

➾ para todo $n \geq 1$. (Verifique você mesmo.)

Material com direitos autorais



























Apêndice D *Polinômios*

Euler deu a mais algébrica das demonstrações
da existência das raízes de uma equação polinomial...
Acho injustiça atribuir tal prova exclusivamente a Gauss,
que meramente adicionou os retoques finais.

Georg Frobenius, 1907
Citado no arquivo MacTutor History of Mathematics,
http://www-history.mcs.st-and.ac.uk/history/

Um ***polinômio*** é uma função p de uma única variável x que pode ser escrita na forma

$$p(x) = a_0 + a_1x + a_2x^2 + \cdots + a_nx^n \tag{1}$$

onde $a_0, a_1, \ldots, a_n$ são constantes ($a_n \neq 0$), chamadas ***coeficientes*** de p.
Convencionando que $x^0 = 1$, podemos usar o símbolo de somatório para escrever p como

$$p(x) = \sum_{k=0}^{n} a_k x^k$$

O inteiro n chama-se ***grau*** de p, e é denotado como gr $p = n$. Um polinômio de grau zero é chamado de ***polinômio constante***.

EXEMPLO 1 Quais das funções a seguir são polinômios?

(a) $2 - \frac{1}{3}x + \sqrt{2}x^2$ (b) $2 - \dfrac{1}{3x^2}$ (c) $\sqrt{2x^2}$

(d) $\ln\left(\dfrac{2e^{5x^3}}{e^{3x}}\right)$ (e) $\dfrac{x^2 - 5x + 6}{x - 2}$ (f) $\sqrt{x}$

(g) $\cos(2\cos^{-1}x)$ (h) e^x

SOLUÇÃO: (a) Esta é a única expressão que se percebe facilmente ser um polinômio.

(b) Um polinômio da forma mostrada em (1) não pode ter limite infinito quando x se aproxima de um valor finito $[\lim_{x \to c} p(x) \neq \pm\infty]$, ao passo que $2 - 1/3x^2$ se aproxima de $-\infty$ quando x se aproxima de zero. Portanto, a expressão em (b) não é um polinômio.

(c) Temos

$$\sqrt{2x^2} = \sqrt{2}\sqrt{x^2} = \sqrt{2}|x|$$

Material com direitos autorais









Observação • O Teorema da Fatoração estabelece uma relação entre as raízes de um polinômio e seus fatores *lineares*. No entanto, um polinômio sem fatores lineares pode ter fatores de grau maior. Além disso, quando fatoramos um polinômio, precisamos levar em conta o sistema numérico no qual os fatores estão sendo considerados.

Por exemplo, considere o polinômio

$$p(x) = x^4 + 1$$

Pelo Teorema das Raízes Racionais, as únicas raízes possíveis de p sobre o conjunto $\mathbb{Q}$ dos *números racionais* são 1 e –1. Uma rápida verificação mostra que nenhuma delas funciona; portanto, $p(x)$ não possui nenhum fator linear com coeficientes racionais. Vamos procurar os fatores quadráticos usando o método dos ***coeficientes a determinar***.

Suponhamos que

$$x^4 + 1 = (x^2 + ax + b)(x^2 + cx + d)$$

Efetuando a multiplicação do lado direito e comparando os coeficientes, obtemos as seguintes equações:

$$\begin{aligned} a + c &= 0 \\ b + ac + d &= 0 \\ bc + ad &= 0 \\ bd &= 1 \end{aligned}$$

Se $a = 0$, então $c = 0$ e $d = -b$. Isso dá $-b^2 = 1$, que não possui solução em $\mathbb{Q}$. Portanto, vamos assumir que $a \neq 0$. Então, $c = -a$, e obtemos $d = b$. Daí segue que $b^2 = 1$; logo, $b = 1$ ou $b = -1$. Isso implica que $a^2 = 2$ ou $a^2 = -2$, respectivamente, e nenhuma delas tem solução em $\mathbb{Q}$. Concluímos que $x^4 + 1$ não pode ser fatorado em $\mathbb{Q}$. Dizemos que ele é ***irredutível*** sobre $\mathbb{Q}$.

Entretanto, $x^4 + 1$ se fatora sobre o conjunto $\mathbb{R}$ dos *números reais*. Os cálculos que fizemos mostram justamente que

$$x^4 + 1 = (x^2 + \sqrt{2}x + 1)(x^2 - \sqrt{2}x + 1)$$

(Por quê?) Para verificar se é possível fatorar mais, aplicamos a fórmula quadrática. O primeiro fator possui os seguintes zeros:

$$x = \frac{-\sqrt{2} \pm \sqrt{(\sqrt{2})^2 - 4}}{2} = \frac{-\sqrt{2} \pm \sqrt{-2}}{2} = \frac{\sqrt{2}}{2}(-1 \pm i) = -\frac{1}{\sqrt{2}} \pm \frac{1}{\sqrt{2}}i$$

que estão em $\mathbb{C}$ mas não em $\mathbb{R}$. Com isso, $x^2 + \sqrt{2}x + 1$ não pode ser escrito como produto de fatores lineares em $\mathbb{R}$. Analogamente, $x^2 - \sqrt{2}x + 1$ não pode ser escrito como produto de fatores lineares em $\mathbb{R}$.

Nossos cálculos mostram que é possível uma fatoração completa de $x^4 + 1$ sobre o conjunto $\mathbb{C}$ dos *números complexos*. As quatro raízes de $x^4 + 1$ são

$$\alpha = -\frac{1}{\sqrt{2}} + \frac{1}{\sqrt{2}}i, \quad \overline{\alpha} = -\frac{1}{\sqrt{2}} - \frac{1}{\sqrt{2}}i, \quad -\overline{\alpha} = \frac{1}{\sqrt{2}} + \frac{1}{\sqrt{2}}i, \ -\alpha = \frac{1}{\sqrt{2}} - \frac{1}{\sqrt{2}}i$$

que, como mostra a Figura 1, estão todas na circunferência unitária no plano complexo. Assim, a fatoração de $x^4 + 1$ é

$$x^4 + 1 = (x - \alpha)(x - \overline{\alpha})(x + \overline{\alpha})(x + \alpha)$$

Material com direitos autorais





You have either reached a page that is unavailable for viewing or reached your viewing limit for this book.































A sintaxe seguinte faz a mesma coisa que a anterior:

```
»A\b
ans =

   3.0000

  -1.0000

   2.0000
```

Você também pode obter a solução de um sistema a partir da forma escalonada reduzida por linhas da sua matriz completa. (Isso é necessário se o sistema possui infinitas soluções ou nenhuma, já que o Matlab não pode tratar esses casos.)

```
»Ab=[1 -1 -1 2; 3 -3 2 16; 2 -1 1 9];
»rref(Ab)
ans =

  1    0    0    3

  0    1    0   -1

  0    0    1    2
```

Caso prefira, você pode pedir ao Matlab para passar por todo o processo de redução por linhas. (Tenha em mente que o Matlab usa estratégias de pivoteamento parcial.) Isso é feito com o comando

```
rrefmovie(Ab)
```

Use o comando `mod` em Matlab para tratar sistemas sobre $\mathbb{Z}_p$.

Exemplo 10, Seção 2.3

```
»A=[1 2 1; 1 0 1; 0 1 2]; b=[0; 2; 1];
»mod(A\b, 3)
ans =

  1

  2

  1
```

Essa abordagem, entretanto, às vezes precisa de um pouco de ajuda. Suponha que você precise resolver o sistema $A\mathbf{x} = b$, onde

$$A = \begin{bmatrix} 1 & 1 & 2 \\ 3 & 0 & 1 \\ 2 & 2 & 3 \end{bmatrix} \quad \text{e} \quad b = \begin{bmatrix} 3 \\ 3 \\ 2 \end{bmatrix}$$

sobre $\mathbb{Z}_5$. Você obtém

```
»A=[1 1 2;3 0 1;2 2 3]; b=[3;3;2];
»sol=mod(A\b,5)
sol =

  14/3

  1/3

  4
```

Isso não é o que você quer, já que a solução não é um vetor com elementos em $\mathbb{Z}_5$. Nesse caso, você precisa usar o seguinte teorema, de Euler:

Material com direitos autorais







O Matlab possui um comando predefinido para a fatoração *LU* de uma matriz.

```
»A=[2 1 3; 4 -1 3; -2 5 5];
»[L, U] = lu(A)
L =

      0.5000    0.3333    1.0000
      1.0000         0         0
     -0.5000    1.0000         0

U =

      4.0000   -1.0000    3.0000
           0    4.5000    6.5000
           0         0   -0.6667
```

Para verificar que esse resultado está correto, calcule

```
»L*U
ans =

      2.0000    1.0000    3.0000
      4.0000   -1.0000    3.0000
     -2.0000    5.0000    5.0000
```

que é *A*, como esperado.

Capítulo 4: Autovalores e Autovetores

O Matlab pode tratar autovalores, autovetores e polinômios característicos.

Exemplo 1, Seção 4.4

```
»A=[0 1 0; 0 0 1; 2 -5 4];
```

O polinômio característico de *A* pode ser obtido pelo comando

```
»c=poly(A)
c =

     1.0000   -4.0000    5.0000   -2.0000
```

Esse comando fornece os coeficientes do polinômio característico, ordenado de forma crescente pelo grau de x. Para ver o polinômio de fato, converta esse resultado para a forma padrão:

```
»poly2sym(c)
ans =

x^3-4*x^2+5*x-2
```

Para resolver a equação característica (para obter os autovalores de *A*), calcule as raízes de *c*.

```
»roots(c)
ans =

     2.0000
     1.0000 + 0.0000i
     1.0000 - 0.0000i
```

Material com direitos autorais







Exercícios 2.6

1.

n	0	1	2	3	4	5
x_1	0	0,8571	0,9714	0,9959	0,9991	0,9998
x_2	0	0,8000	0,9714	0,9943	0,9992	0,9998

Solução exata: $x_1 = 1$ e $x_2 = 1$

3.

n	0	1	2	3	4	5	6
x_1	0	0,2222	0,2539	0,2610	0,2620	0,2622	0,2623
x_2	0	0,2857	0,3492	0,3582	0,3603	0,3606	0,3606

Solução exata (com quatro casas decimais): $x_1 = 0{,}2623$ e $x_2 = 0{,}3606$

5.

n	0	1	2	3	4	5	6	7	8
x_1	0	0,3333	0,2500	0,3055	0,2916	0,3009	0,2986	0,3001	0,2997
x_2	0	0,2500	0,0834	0,1250	0,0972	0,1042	0,0996	0,1008	0,1000
x_3	0	0,3333	0,2500	0,3055	0,2916	0,3009	0,2986	0,3001	0,2997

Solução exata: $x_1 = 0{,}3$, $x_2 = 0{,}1$ e $x_3 = 0{,}3$

7.

n	0	1	2	3	4
x_1	0	0,8571	0,9959	0,9998	1,0000
x_2	0	0,9714	0,9992	1,0000	1,0000

O uso do método de Gauss-Seidel proporciona uma margem de erro de 0,001 da resposta exata, após três iterações. Para obter a mesma acuidade usando o método de Jacobi, precisamos fazer quatro iterações.

9.

n	0	1	2	3	4
x_1	0	0,2222	0,2610	0,2622	0,2623
x_2	0	0,3492	0,3603	0,3606	0,3606

O uso do método de Gauss-Seidel proporciona uma margem de erro de 0,001 da resposta exata, após três iterações. Para obter a mesma acuidade usando o método de Jacobi, precisamos fazer quatro iterações.

Material com direitos autorais







ÁLGEBRA LINEAR

Este livro foi estruturado de forma bastante flexível, com a preocupação central de que a álgebra linear constitua um assunto ainda mais estimulante e de fácil aprendizado, tornando-a mais acessível ao estudante dessa matéria. Escrito de forma clara, direta e objetiva, aborda temas como *Vetores, Matrizes, Autovalores e Autovetores, Ortogonalidade, Espaços Vetoriais, e Distância e Aproximação.*

A apresentação de conceitos-chave com antecedência, a ênfase em vetores e geometria e os inúmeros exercícios e exemplos que reforçam o fato de a álgebra linear ser uma ferramenta valiosa para a modelagem de problemas da vida real consistem no principal diferencial deste livro. A apresentação de pequenos esboços biográficos de muitos dos matemáticos que contribuíram para o desenvolvimento da álgebra linear é outro diferencial, valorizando a história da matemática.

Aplicações

Destina-se à disciplina introdução à álgebra linear, nos cursos de Engenharia, Física, Química, Ciência da Computação, Matemática, Estatística, Economia, Administração de Empresas, entre outros.

ISBN 85-221-0359-3

9 788522 103591

THOMSON™

Visite nossos sites:
www.thomsonlearning.com.br
www.thomsonlearning.com

Material com direitos autorais